THE PROCEEDINGS OF THE PHYSICAL SOCIETY

VOL. 60, PART 4 — 1 April 1948 — No. 340

The Electronic Structure and Bond Lengths of Coronene and Pyrene

BY W. E. MOFFITT AND C. A. COULSON*

Physical Chemistry Laboratory, Oxford

* Now at King's College, London

MS. received 1 *September* 1947

ABSTRACT. Molecular orbital and valence-bond pairing treatments are given for the bond lengths of coronene and pyrene. In both cases the agreement of the molecular orbital treatment with the experimental values is good, and in both cases it is rather better than the agreement of the valence-bond method, though certain characteristic features are common to both treatments and are found experimentally. The Fries rule, which governs the relative importance of the various Kekulé structures in valence-bond theory, is shown to be only partially applicable. A more valid rule is suggested for condensed hydrocarbons generally.

§1. INTRODUCTION

THE recent very accurate X-ray work of Robertson and White (1945, 1947) on the dimensions of the polynuclear hydrocarbons coronene and pyrene is of great interest. It discloses a new testing ground for the various theories of the electronic structure of these aromatic systems. In this paper we give the results of two alternative theoretical treatments of these molecules. In the first, using the molecular orbital method of Hund, Mulliken, Hückel and Lennard-Jones, we express the electronic wave function as a product of non-localized one-electron molecular orbitals (called MO), constructed as linear combinations of atomic orbitals (LCAO). In the second (HLSP) method, due to Heitler, London, Slater and Pauling, the total wave function is obtained as a linear combination of many-electron functions or structures, in each of which the electrons are perfectly paired; since these latter calculations would be extremely tedious if followed to the full, we examine an approximation to them.

§2. THE MOLECULAR ORBITAL METHOD

The characteristic properties of aromatic systems are attributed to those electrons in orbitals which are antisymmetrical with respect to reflections in the planes of these molecules. We shall consider the behaviour of these (π) unsaturation, or mobile, electrons alone, and make the usual assumption that their interaction with the remaining σ electrons is negligible for our purposes. The LCAO MO procedure has been fully described elsewhere (e.g. Lennard-Jones and Coulson 1939), so that we shall merely quote the results of our calculations for pyrene. The treatment for coronene is of greater interest, however, owing to the high symmetry of the molecule.

The symmetry elements of coronene, like those of benzene, belong to the point group D_{6h}. The 24 $2p_z$ atomic orbitals, one for each carbon atom (figure 1), which we call the $\phi_i (i=1, 2, \ldots 24)$, constitute a reducible representation Γ of this point group. The characters of Γ are found to be

$$\chi(E)=24,\ \chi(C_2)=\chi(C_3)=\chi(C_6)=\chi(C_2'')=0,\ \chi(C_2')=-4,$$
$$\chi(\sigma_h)=-24,\ \chi(i)=\chi(S_6)=\chi(S_3)=\chi(\sigma_v'')=0,\ \chi(\sigma_v')=-4.$$

The notation is fully explained by Wilson (1934). Accordingly, by the theory of group characters, Γ has irreducible components

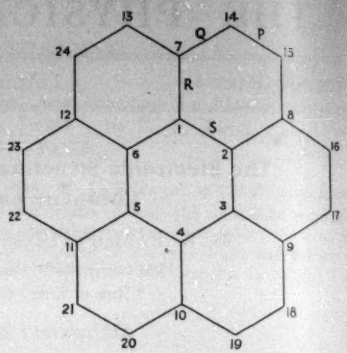


Figure 1. Coronene $C_{24}H_{12}$.

$$\Gamma = A_{1u} + 3A_{2u} + B_{1g} + 3B_{2g} + 4E_{1g} + 4E_{2u}. \qquad \ldots\ldots(1)$$

Following a method analogous to that used by Bloch (1929) in his discussion of metallic structure, we define the functions

$$\left.\begin{aligned}
\kappa(\omega) &= \phi_1 + \omega\phi_2 + \omega^2\phi_3 + \omega^3\phi_4 + \omega^4\phi_5 + \omega^5\phi_6,\\
\lambda(\omega) &= \phi_7 + \omega\phi_8 + \omega^2\phi_9 + \omega^3\phi_{10} + \omega^4\phi_{11} + \omega^5\phi_{12},\\
\mu(\omega) &= \phi_{13} + \omega\phi_{15} + \omega^2\phi_{17} + \omega^3\phi_{19} + \omega^4\phi_{21} + \omega^5\phi_{23},\\
\nu(\omega) &= \phi_{14} + \omega\phi_{16} + \omega^2\phi_{18} + \omega^3\phi_{20} + \omega^4\phi_{22} + \omega^5\phi_{24}.
\end{aligned}\right\} \qquad \ldots\ldots(2)$$

The numbering system of the ϕ_i is as in figure 1, and ω is a number, in general complex, of unit modulus. It may now readily be shown that the required LCAO forms for the molecular orbitals are

$$\left.\begin{aligned}
&A_{1u}:\ \mu(1)-\nu(1),\\
&A_{2u}:\ k\kappa(1)+l\lambda(1)+m\{\mu(1)+\nu(1)\},\\
&B_{1g}:\ \mu(-1)+\nu(-1),\\
&B_{2g}:\ k'\kappa(-1)+l'\lambda(-1)+m'\{\mu(-1)-\nu(-1)\},\\
&E_{1g}:\ k''\kappa(\epsilon)+l''\lambda(\epsilon)+m''\mu(\epsilon)+n''\nu(\epsilon)\\
&E_{2u}:\ k'''\kappa(\epsilon^2)+l'''\lambda(\epsilon^2)+m'''\mu(\epsilon^2)+n'''\nu(\epsilon^2),
\end{aligned}\right\} \quad \epsilon = e^{2\pi i/3} \text{ or } e^{5\pi i/3}. \qquad \ldots\ldots(3)$$

In these formulae $k, l, m, k', \ldots, n'''$ are constants, one set being found from each solution of the secular equation. These sets of coefficients are easily calculated, for, according to (1), the secular equation is a cubic for the levels of symmetry A_{2u} and B_{2g}, it is linear for A_{1u} and B_{1g}, and a biquadratic for the degenerate levels E_{1g}, E_{2u}. The secular equations may be written down when the matrix components of $\mathscr{H}-E$ in the representation (2) are known, where $\mathscr{H}$ is the Hamiltonian and E is the energy, as usual. We have

$$(\kappa|\mathscr{H}-E|\kappa)=6x\beta+6(\omega+\omega^5)\beta,$$
$$(\kappa|\mathscr{H}-E|\lambda)=(\lambda|\mathscr{H}-E|\mu)=(\lambda|\mathscr{H}-E|\nu)=6\beta,$$
$$(\lambda|\mathscr{H}-E|\lambda)=(\mu|\mathscr{H}-E|\mu)=(\nu|\mathscr{H}-E|\nu)=6x\beta,$$
$$(\mu|\mathscr{H}-E|\nu)=(\nu|\mathscr{H}-E|\mu)^*=6\omega\beta,$$

where β is the resonance integral between adjacent $2p_z$ orbitals, E_0 is the atomic energy term (Coulomb term) defined by $E_0=(\phi_l|\mathscr{H}|\phi_l)$, and $x=(E_0-E)/\beta$. The secular equations thus reduce to

$$A_{1u}:\ 0=x-1,$$

$$A_{2u}:\ 0=\begin{vmatrix} x+2 & 1 & 0 \\ 1 & x & 2 \\ 0 & 2 & 2x+2 \end{vmatrix} \equiv \Delta(x),$$

$$B_{1g}:\ 0=x+1,$$

$$B_{2g}:\ 0=\begin{vmatrix} x-2 & 1 & 0 \\ 1 & x & 2 \\ 0 & 2 & 2x-2 \end{vmatrix} \equiv -\Delta(-x),$$

$$E_{1g}:\ 0=\begin{vmatrix} x+1 & 1 & 0 & 0 \\ 1 & x & 1 & 1 \\ 0 & 1 & x & \epsilon \\ 0 & 1 & \epsilon^* & x \end{vmatrix} \equiv -\tfrac{1}{2}(x+1)\Delta(-x-1),$$

$$E_{2u}:\ 0=\begin{vmatrix} x-1 & 1 & 0 & 0 \\ 1 & x & 1 & 1 \\ 0 & 1 & x & \epsilon^2 \\ 0 & 1 & \epsilon^{*2} & x \end{vmatrix} \equiv \tfrac{1}{2}(x\quad 1)\Delta(x\quad 1).$$

Thus the roots are all either ± 1 or simply expressible in terms of the roots of the cubic equation for the levels A_{2u}. The resultant spectrum of energy levels may be represented graphically as in the accompanying diagram (figure 2). In the ground state we fill the orbitals of lowest energy with electrons, two at a time, in accordance with the Pauli exclusion principle. These electrons, whose spins are anti-parallel in pairs, are represented by crosses on our diagram. The final electronic configuration may thus be written

$$^1A_{1g}: (a_{2u})^2(e_{1g})^4(e_{2u})^4(\overline{a_{2u}})^2(b_{2g})^2(b_{1g})^2(\overline{e_{1g}})^4(\overline{e_{2u}})^4.$$

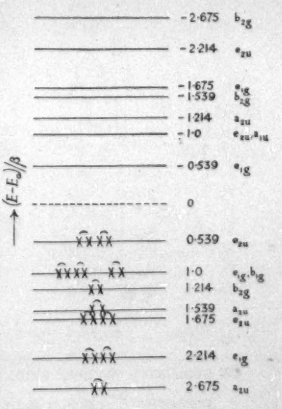


Figure 2. Energies of molecular orbitals in coronene

The explicit LCAO forms for each MO are readily calculated at this stage, since the solutions of the secular equations enable us to determine the values of the sets of coefficients k, l, m, n in (3). That is, each MO may be written in the form

$$\Psi^{(r)}=\sum_i c_i^{(r)}\phi_i \quad (r, i=1, 2, \ldots, 24),$$

where the $c_i^{(r)}$ are known, and satisfy the normalizing condition $\sum_i |c_i^{(r)}|^2=1$. The mobile (or π) bond order p_{ij} of a given bond between atoms i and j may now be calculated, for it is the sum of separate contributions from each of the various π electrons. Following Coulson (1938), the calculation

$$p_{ij}=\sum_r p_{ij}^{(r)}=\sum \tfrac{1}{2}(c_i^{(r)}c_j^{(r)*}+c_i^{(r)*}c_j^{(r)})$$

is readily accomplished. The final values have been tabulated under column 1 for coronene (table 1) and pyrene (table 2). The internuclear distances which these would lead us to predict are given in column 4. Part, but not all, of the information in table 1 has already been given in a preliminary note (Coulson 1944).

§ 3. THE PAIRING METHOD

The HLSP, valence-bond, or pairing, method of treating aromatic molecules consists in selecting that linear combination of the structures ψ_t which gives the minimum energy to the particular molecule under consideration. The procedure is well known and has been applied to many molecules by Pauling (1933) and others. The complete treatment of our molecules along these lines would be an extremely arduous task. For coronene we should have to take more than 200,000 non-ionic structures into account. Therefore most of these structures are neglected, and only a reasonable number are considered to contribute in any marked extent to the ground states of these molecules. It is customary, especially in qualitative discussion, to make the simplification of neglecting all ionic and excited structures containing formal bonds (non-adjacent pairing). Accordingly we confine ourselves to the remaining non-excited, or Kekulé, structures (figure 3), which number 20 and 6 for coronene and pyrene respectively.

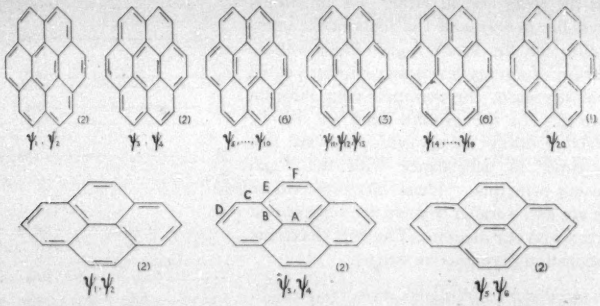


Figure 3. Canonical structures for coronene (top line) and pyrene (bottom line).

By symmetry we may write the two wave functions of the ground states as:

coronene $\Psi = a(\psi_1 + \psi_2) + b(\psi_3 + \psi_4) + c(\psi_5 + \ldots + \psi_{10}) + d(\psi_{11} + \psi_{12} + \psi_{13}) + e(\psi_{14} + \ldots + \psi_{19}) + f\psi_{20}$,

pyrene $\Psi = r(\psi_1 + \psi_2) + s(\psi_3 + \psi_4) + t(\psi_5 + \psi_6)$.

The secular equations are of degree 6 and 3 respectively. They may be set up with the help of the symbolic technique devised by Pauling (1933), and they are easily solved. For the two ground states

$E\,(\text{coronene}) = Q + 6{\cdot}64J$, and $a:b:c:d:e:f = 0{\cdot}50:0{\cdot}60:0{\cdot}76:0{\cdot}75:0{\cdot}88:1{\cdot}00$,

$E\,(\text{pyrene}) = Q + 4{\cdot}70J$, and $r:s:t = 1{\cdot}31:1{\cdot}27:1{\cdot}00$.

This corresponds to resonance energies $3{\cdot}64J$ and $2{\cdot}20J$. Q and J are the Coulomb and exchange integrals as usual.

The bond orders may now be defined in one of two alternative ways. The first is due to Pauling, Brockway and Beach (1935), which we shall abbreviate as (PBB). The second is due to Penney (1937), using a technique devised by Dirac; we shall abbreviate this as (PD). The latter definition of p_{ij} as $p_{ij} = -(\frac{4}{3})\ \overline{\mathbf{s}_i \cdot \mathbf{s}_j}$, where $\mathbf{s}_i$ and $\mathbf{s}_j$ are the spin matrices of the electrons on atoms i and j, is fundamentally more satisfactory than the earlier one which is merely the weighted mean of double-bond character as revealed by the weights of the various structures. The (PD) bond orders are shown in column 2 of tables 1 and 2; the (PBB) bond orders are shown in column 3. In columns 5 and 6 we give the (PD) and the (PBB) predictions of the bond distances (in A.), using appropriate curves of order against length in each case. The rows labelled $\bar{p}$ and $\bar{r}$ in the tables give the average bond order and internuclear distances respectively.

Table 1. Coronene

	1	2	3	4	5	6	7
	MO (Coulson)	Pairing (PD)	Pairing (PBB)	MO (Coulson)	Pairing (PD)	Pairing (PBB)	Observed (Robertson and White)
P	0·757	0·816	0·742	1·372	1·362	1·362	1·385
Q	0·538	0·316	0·258	1·411	1·450	1·438	1·415
R	0·538	0·629	0·484	1·411	1·387	1·393	1·43
S	0·522	0·328	0·258	1·415	1·446	1·438	1·43
$\bar{p}$, $\bar{r}$	0·576	0·481	0·400	1·405	1·419	1·414	1·415

Table 2. Pyrene

	1	2	3	4	5	6	7
	MO (Coulson)	Pairing (PD)	Pairing (PBB)	MO (Coulson)	Pairing (PD)	Pairing (PBB)	Observed (Robertson and White)
A	0·536	0·280	0·231	1·412	1·458	1·445	1·45
B	0·524	0·516	0·384	1·415	1·408	1·410	1·39
C	0·594	0·587	0·500	1·401	1·377	1·390	1·42
D	0·669	0·587	0·500	1·388	1·377	1·390	1·39
E	0·503	0·141	0·116	1·420	1·495	1·484	1·45
F	0·777	0·931	0·884	1·370	1·347	1·348	1·39
$\bar{p}$, $\bar{r}$	0·592	0·498	0·421	1·403	1·409	1·412	1·41

Columns 1–3 are bond orders, columns 4–7 are bond lengths.
The last decimal in columns 4, 5 and 6 for the bond lengths (A.) is unreliable on an absolute scale, but is useful for comparisons.

§4. DISCUSSION

Robertson and White (1945, 1947) have determined the dimensions of coronene and pyrene by means of X-ray crystallography, and we have reproduced their values in tables 1 and 2. First, we note that for coronene the agreement between the LCAO MO predictions and the empirical data is very close indeed. The average discrepancy is only 0·01 A., and in no case is it more than 0·02 A.—quite within the limits of accuracy of the X-ray determinations. The pairing method gives somewhat poorer results, although the average internuclear distance agrees very well with the measured value. But there are relatively large variations in the lengths of individual bonds, which are apparently not found experimentally. Thus both the X-ray and MO methods gives the bonds R and S a practically equal value, whereas the pairing method expects these to be separated by some

0·05 or 0·06 A. As Robertson and White have remarked, using only the non-ionic non-excited structures, the empirical data can be reconciled with this representation if the relative weights of the individual structures follow a scheme which is the exact opposite of that which we have calculated; the failure of this set of structures to distinguish between the bonds Q and S is also to be remarked.

It has been suggested (the Fries rule) that those Kekulé structures which show the greatest number of benzenoid rings should have the greatest weight in the superposition diagram. In table 3 we show that this is roughly valid for coronene. Certainly the greatest weight is for ψ_{20}, which has the greatest number (6) of benzenoid rings; and the least weight is for ψ_1 and ψ_2 with the smallest number (1). But apart from these extremes, the weights do not support the rule, which should therefore only be granted a relative validity. A better rule seems to be that those structures which show the greatest number of double bonds in the "exposed" positions (e.g. P of coronene) have the greatest weight. Table 3 shows that there are effectively no exceptions to this in coronene. It seems as if there was a tendency for such outer bonds, in which both ends of the bond are only attached to one other aromatic carbon, to be double; both phenanthrene and pyrene show the same characteristic. For pyrene, shown in table 4, the two types of structure with most exposed bonds (here only 2) have a similar weight, which is greater than that of the remaining structures with only one exposed bond. But the situation regarding the numbers of benzenoid rings is much the same as in coronene.

Table 3. Coronene

Structures	Coefficients	N_b	N_e
ψ_1, ψ_2	0·501	1	3
ψ_3, ψ_4	0·603	4	3
$\psi_5 \cdots \psi_{10}$	0·757	3	4
ψ_{11}, ψ_{12}, ψ_{13}	0·749	2	4
$\psi_{14} \cdots \psi_{19}$	0·879	4	5
ψ_{20}	1·000	6	6

Table 4. Pyrene

Structures	Coefficients	N_b	N_e
ψ_1, ψ_2	1·31	2	2
ψ_3, ψ_4	1·27	3	2
ψ_5, ψ_6	1·00	1	1

N_b=number of benzenoid rings ; N_e=number of exposed bonds.

A similar comparison between theoretical and experimental bond lengths may be made for pyrene as for coronene. We notice, however, that the difference between the MO and x-ray values is larger here than in coronene. This is possibly due to a lack of certainty in the x-ray results, since only nine of the sixteen carbons are "seen" in the diffraction pattern of the crystal. Again the LCAO MO method gives the more likely of the theoretical values. The high double bond character of the exposed bond F is, however, expected by all these methods—in agreement with the chemistry of this molecule.

The relatively poorer values given by the HLSP treatment are readily attributed to the lack of a truly representative set of structures. In conjunction with the findings of Daudel and Pullman (1946) for a rather different set of polynuclear hydrocarbons, we may therefore conclude that in these condensed-ring molecules the contribution of excited structures to the ground states is considerable, since the LCAO MO and the full HLSP calculations using only non-ionic structures generally lead to very similar results for those simpler systems where they can be carried through completely.

REFERENCES

BLOCH, 1929, *Z. Phys.*, **52**, 555.
COULSON, 1938, *Proc. Roy. Soc.* A, **169**, 413.
COULSON, 1944, *Nature, Lond.*, **154**, 797.
DAUDEL and PULLMAN, 1946, *J. Phys. et Radium*, **7**, 105.
LENNARD-JONES and COULSON, 1939, *Trans. Faraday Soc.*, **35**, 811.
PAULING, 1933, *J. Chem. Phys.*, **1**, 280.
PAULING, BROCKWAY and BEACH, 1935, *J. Amer. Chem. Soc.*, **57**, 2705.
PENNEY, 1937, *Proc. Roy. Soc.* A, **158**, 306.
ROBERTSON and WHITE, 1945, *J. Chem. Soc.*, 607 ; 1947, *Ibid.*, 358.
WILSON, 1934, *Phys. Rev.*, **45**, 706.

The Influence of Surface Films on the Electrical Behaviour of Contacts

BY (MISS) C. C. DILWORTH
H. H. Wills Physical Laboratory, University of Bristol

Communicated by N. F. Mott; MS. received 24 June 1947

ABSTRACT. The variation of current with voltage at an idealized contact between two crystals of a semiconductor is calculated on the assumption that electrons penetrate the surface barrier by tunnel effect. Comparison with experimental curves for silicon carbide powders leads to the conclusion that these crystals are covered by an insulating surface film. The existence of such a film affects the rectifying properties of the crystal when it is in contact with a metal. It is shown that this can account for the discrepancies observed between experimental curves and those deduced from the simple Schottky theory of rectification.

§1. INTRODUCTION

THE non-ohmic resistance between a metal and a semiconductor can be due either to the existence of a contact potential between them (Schottky's theory) or to the existence of an insulating layer of different composition (Wilson's theory). In this paper we examine the effect of an insulating layer present in addition to the Schottky layer, the current penetrating the insulating layer by tunnel effect. This model seems correct for silicon carbide with thin quartz layers on the surface. Rectification is in general in the same direction as in the Schottky theory, but the rectifying properties are less good. We consider also the effect of these layers on the contacts between particles of a silicon carbide powder.

§2. THE CONTACT BETWEEN TWO SEMICONDUCTORS

The resistance of the contact between two crystals of silicon carbide is of the order of 10^{10} times the resistance of the main bulk of the crystals for potential differences of less than 10^{-2} volt. As the voltage is increased, the resistance falls rapidly, as shown by figure 1, reproduced from the paper by Braun and Busch (1942). We conclude from this and earlier work (Kurtschatow *et al.* 1935) that

this is due to the existence of an insulating film on the surface, and show that it is impossible to explain the magnitude of the effect otherwise.

As a model for calculation, the contact is represented by two spheres, of radius r, in point contact. The area of contact of the surfaces is then zero, but electrons can pass through the air-space in a narrow zone surrounding the point of contact. For a given value of the angle θ between the common axis OX and the radius vector r, the width of the interspace measured in the direction of the common axis is (figure 2):

$$x = 2r(1 - \cos\theta). \qquad \ldots\ldots(1)$$

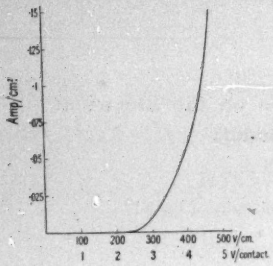


Figure 1. The current-voltage characteristic for silicon carbide crystal contacts.

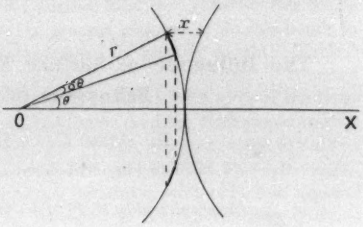


Figure 2. Idealized contact between two spheres.

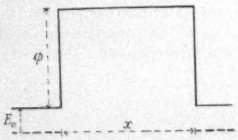


Figure 3. The potential energy of an electron in the interspace.

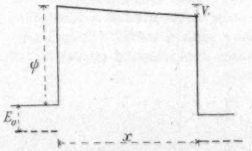


Figure 4. The potential barrier in an external field.

The potential V of an electron in the interspace is represented in figure 3. E_0 is the activation energy of the semiconductor; thus the conductivity depends on temperature according to the formula

$$\sigma = \sigma_0 e^{-E_0/2kT}. \qquad \ldots\ldots(2)$$

$e\phi$ is the work function of the semiconductor; the fraction of incident electrons having sufficient energy to surmount the potential barrier is then

$$f = e^{-e\phi/kT}. \qquad \ldots\ldots(3)$$

Since $e\phi$ is estimated to be of the order of 2–4 ev., this is negligible at room temperatures ($kT \sim 2{\cdot}5 \times 10^{-2}$ ev.). The chief mechanism by which electrons cross the interspace is, therefore, that of "tunnel effect".

On this mechanism the current may be calculated as follows. Under an applied voltage V the potential barrier takes the form shown in figure 4.

The finite probability, P, that an electron with insufficient energy to surmount the barrier should penetrate it is given by

$$P \sim e^{-\alpha x/V}, \qquad \ldots\ldots(4)$$

where

$$\alpha = \tfrac{4}{3}\eta[(\phi - V)^{3/2} - \phi^{3/2}], \qquad \eta = 2\boldsymbol{m}/\hbar,$$

and $\boldsymbol{h} = 2\pi\hbar$ is Planck's constant.

The number of electrons incident on unit area of the barrier per second from left to right is

$$N_1 = \tfrac{1}{2}N\sqrt{(2\boldsymbol{k}T/\pi\boldsymbol{m})}. \qquad \ldots\ldots(5)$$

From right to left the number with energy great enough to enter the second conductor is

$$N_2 = \tfrac{1}{2}N\sqrt{(2\boldsymbol{k}T/\pi\boldsymbol{m})}e^{-\boldsymbol{e}V/\boldsymbol{k}T}, \qquad \ldots\ldots(6)$$

where N is the number of free electrons in unit volume of the semiconductor. The element of area, normal to the field in the angular range $\theta_0 < \theta < \theta_0 + d\theta$, is

$$dA = 2\pi r^2 \sin\theta_0\, d\theta \cos\theta_0. \qquad \ldots\ldots(7)$$

The current crossing this area is

$$\begin{aligned} dI &= \boldsymbol{e}(N_1 - N_2)P\,dA \\ &= N\{\sqrt{(\boldsymbol{k}T/2\pi\boldsymbol{m})}\}(1 - e^{-\boldsymbol{e}V/\boldsymbol{k}T})e^{-\alpha \,.\, 2r(1-\cos\theta)/V} \,.\, 2\pi r^2 \sin\theta_0 \cos\theta_0\, d\theta. \qquad \ldots\ldots(8) \end{aligned}$$

Integrating over θ the total current crossing the contact is

$$I = \int_{\theta=0}^{\theta=\frac{1}{2}\pi} dI = N\{\sqrt{(\boldsymbol{k}T/2\pi\boldsymbol{m})}\}(1 - e^{-\boldsymbol{e}V/\boldsymbol{k}T})2\pi r^2 f(\alpha r/V), \qquad \ldots\ldots(9)$$

where

$$f\left(\frac{\alpha r}{V}\right) = \frac{e^{-\alpha r/V}}{(\alpha r/V)^2} - \frac{1}{(\alpha r/V)^2} + \frac{1}{(\alpha r/V)}.$$

If $r \gg \lambda$, where $\lambda^{-1} = 2\kappa\phi^{1/2}$, then $f(\alpha r/V) \sim V/\alpha r$, and the current is given approximately by

$$I = N\{\sqrt{(\boldsymbol{k}T/2\pi\boldsymbol{m})}\}(1 - e^{-\boldsymbol{e}V/\boldsymbol{k}T})2\pi r V/\alpha. \qquad \ldots\ldots(10)$$

The integration over the surface from $\theta = 0$ to $\theta = \theta_0$ approximates closely to that over the whole hemisphere for values of θ_0, such that

$$2r(1 - \cos\theta_0) \geqslant 10\lambda. \qquad \ldots\ldots(11)$$

This means that the emission of electrons is practically confined to the region $0 < \theta < \theta_0$, and there is an effective area of contact

$$\Delta A = \pi a^2, \qquad \ldots\ldots(12)$$

where

$$a^2 = 2r^2(1 - \cos\theta_0) = 10r\lambda.$$

For such a small area the classical resistance due to the constriction of lines of current flow will be considerable. The contact may be represented by a cylinder of the semiconductor of area πa^2 and length L, where

$$L = (10r\lambda/N)\sqrt{(2\pi\boldsymbol{m}/\boldsymbol{k}T)} \,.\, (\alpha/2\pi)(1 - e^{-\boldsymbol{e}V/\boldsymbol{k}T})^{-1}. \qquad \ldots\ldots(13)$$

The total resistance of each contact is then

$$R=(L+\beta a)/\pi a^2\sigma, \qquad \ldots\ldots(14)$$

where $\beta\sim 0{\cdot}82$ (J. H. Jeans, *Electricity and Magnetism*).

The specific resistance ρ of a mass of crystals of mean diameter $2r_0$, each in contact at a point of average curvature, will be given by

$$\rho=2r_0R=(2r_0/\pi a^2\sigma)(L+\beta a). \qquad \ldots\ldots(15)$$

Curve (*a*) in figure 6 is drawn in the case $r=r_0$, $\phi=4{\cdot}5$ volt and $\sigma=0{\cdot}5$ ohm^{-1} cm^{-1}. For $V<\boldsymbol{kT/e}$ this gives a resistance per contact of the order 10^{-2} ohm, so that it is clear that ρ is of the correct order of magnitude only for $\rho:\rho_0\sim 10^{-8}$, which is not the case. Also the form of equation (10) shows a much slower increase of current with voltage than is actually observed. It therefore appears necessary, in order to obtain a sufficiently high resistance, to assume that the surfaces are separated by an insulating film. The contact is now represented by figure 5.

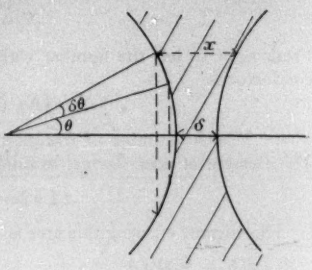


Figure 5. Contact between two semiconductors through a surface film.

The thickness of the barrier

$$x=2r(1-\cos\theta)+\delta_0, \qquad \ldots\ldots(16)$$

where $\frac{1}{2}\delta_0$ is the thickness of the film on each surface. On substituting this value for x in equation (8), the current is given by

$$I=(\sigma/N)\sqrt{(2\pi\boldsymbol{m}/\boldsymbol{kT})}(1-e^{-\boldsymbol{eV/kT}})2\pi r^2 f(\alpha r/V)e^{-\alpha\delta_0}. \qquad \ldots\ldots(17)$$

In the approximate form this is

$$I=(\sigma/N)\sqrt{(2\pi\boldsymbol{m}/\boldsymbol{kT})}(1-e^{-\boldsymbol{eV/kT}})(2\pi rV/\alpha)e^{-\alpha\delta_0}, \qquad \ldots\ldots(18)$$

and the specific resistance is

$$\rho=(2r_0/\sigma)(L'+\beta a)/\pi a^2, \qquad \ldots\ldots(19)$$

where

$$L'=(10r\lambda/N)\sqrt{(\boldsymbol{kT}/2\pi\boldsymbol{m})}(1-e^{-\boldsymbol{eV/kT}})^{-1}e^{\alpha\delta_0}. \qquad \ldots\ldots(20)$$

Curves (*b*), (*c*) and (*d*) in figure 6 correspond to values of $\delta_0=1$, 2 and 10 A. respectively. The resistance for $V<\boldsymbol{kT/e}$ is of the right order for $\delta_0\sim 10$ A. Figure 7 shows how the shape of the theoretical curve compares with experiment. The agreement is not good. Busch's results (Braun and Busch 1942) for rounded grains agree well, but Mitchell's results (Mitchell and Sillars 1947) differ widely from the theoretical curve. Since, however, he finds some rectification even between crystals of the same type, his contacts cannot be of the simple symmetrical form considered and agreement is not to be expected. He does not observe the inflection of the curve at $V\sim\boldsymbol{kT/e}$ due to the factor $(1-e^{-\boldsymbol{eV/kT}})$ and ascribes this to "leakage" through the barrier possibly due to gaps in the oxide layer. It seems more likely to be due to the rectifying properties already noted.

Evidence on the temperature-dependence of the resistance of the barrier is inconclusive. According to tunnel effect mechanism it should be the same as that of the specific resistance of the crystals.

Electron diffraction experiments indicate that there probably exists a surface film of the thickness required.

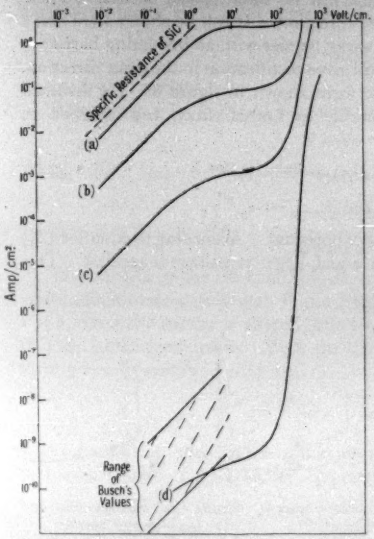


Figure 6. Silicon carbide powder. Theoretical curves.
50 contacts/cm. Work function $\phi = 4{\cdot}5$ v.
Film thickness
(*a*) $\delta = 0$. (*c*) $\delta = 2$ A.
(*b*) $\delta = 1$ A. (*d*) $\delta = 10$ A.

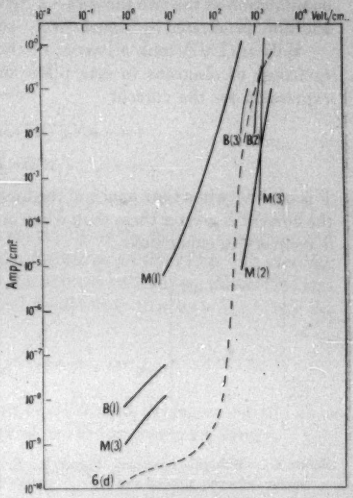


Figure 7. Silicon carbide powder. Experimental curves.
Braun and Busch:
B (1) Region of validity of Ohm's law.
B (2) Rounded grains.
B (3) Angular grains.
Mitchell and Sillars:
M (1) Bonded disc; particle diam. 0·005 cm.
M (2) ,, ,, ,, ,, 0·015 cm.
M (3) ,, ,, ,, ,, 0·025 cm.
6 (*d*) $\delta = 10$ A.; $\phi = 4{\cdot}5$ v. (50 contacts/cm.).

We conclude that the theory that conduction at these contacts takes place by tunnel effect through a surface barrier accounts for the magnitude, and approximately for the variation with voltage, of the resistance of the contact. The problem to be considered next is that of the nature of a contact between such a crystal and a metal.

§3. THE CONTACT BETWEEN A SEMICONDUCTOR AND A METAL

In this section the conduction is always taken to be due to the presence of free electrons. The results can be applied to defect semiconductors, in which positive holes carry the current, by a change in sign of the electronic charge.

The resistance of the contact between a metal and a semiconductor depends not only on the magnitude but also on the direction of the applied voltage. The early theory of rectification, proposed by A. H. Wilson (1932), in which he assumed the existence of a surface barrier penetrated by tunnel effect, gave the wrong sign to the rectification. The more recent theories of Schottky (1939), Mott (1939) and Bethe (1942), which give the right sign, assume that no tunnel effect occurs. These theories give rectification in the observed direction. It will be shown here that a combination of the two theories—a Schottky barrier with an insulating barrier in addition, penetrated by tunnel effect—still gives rectification in the right direction.

Wilson (1937) took a barrier of the form shown in figure 8. He assumed exchange of electrons to take place entirely by tunnel effect, and obtained an expression for the current

$$i = -eN\sqrt{(kT/2\pi m)}\,(e^{eV/kT}-1)e^{-\beta\delta}, \qquad \ldots\ldots(21)$$

$$\beta = 2\eta\phi^{1/2}.$$

V is positive when the metal is at the higher potential. According to equation (21) the current is greater then, than when an equal, negative voltage is applied. This is contrary to experience.

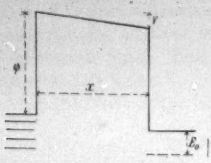


Figure 8. Wilson's potential barrier between a metal and a semiconductor.

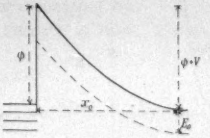


Figure 9. Schottky's space-charge barrier.

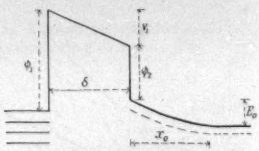


Figure 10. Surface film and space-charge barrier.

Schottky (1939) assumed that the only potential barrier at the contact is that set up within the semiconductor by space-charge, when electrons are lost to the metal by the impurity centres nearest the surface. This type of barrier is shown in figure 9. $e\phi$ is the height of the conduction band of the semiconductor above the Fermi level in the metal, and is given by the difference $e(\phi_1-\phi_2)$ of the work functions of the metal and the semiconductor respectively. In general $e\phi \gg E_0$, so that for simplicity it may be considered that all the impurity centres at distances up to x_0 from the surface are ionized. Then by Poisson's equation, $d^2V/dx^2 = 4\pi\rho/\kappa$,

$$V = (2\pi Ne/\kappa)(2xx_0 - x^2), \qquad \ldots\ldots(22)$$

where the total potential drop

$$V_0 = \phi + V = 2\pi Nex_0^2/\kappa. \qquad \ldots\ldots(23)$$

In order to calculate the current, either "diode theory" (Bethe 1942) or diffusion theory can be used. In the former it is assumed that the barrier is thin compared to the mean free path of the electrons. Then on the average an electron will not suffer collision in passing over the barrier, and the current will be given by

$$i = (\sigma/N)\sqrt{(2\pi m/kT)}\,(1-e^{-eV/kT})e^{-e\phi/kT}. \qquad \ldots\ldots(24)$$

When the barrier thickness is very much greater than the mean free path the diffusion theory must be applied. The current is then obtained from the equation

$$i=\boldsymbol{e}\left\{-\frac{n(x)}{\boldsymbol{kT}}\cdot\frac{dV}{dx}-\frac{dn}{dx}\right\}, \qquad \ldots\ldots(25)$$

where the mobility $v=\boldsymbol{e}D/\boldsymbol{kT}$, and D is the diffusion coefficient. Using the integrating factor $e^{\boldsymbol{e}V/\boldsymbol{kT}}$, this gives for the current,

$$i\int_0^x e^{\boldsymbol{e}V/\boldsymbol{kT}}\,dx=\boldsymbol{e}D\{A-n(x)e^{-\boldsymbol{e}V/\boldsymbol{kT}}\}, \qquad \ldots\ldots(26)$$

where A is a constant.

When $i=0$, $V(x_0)=\phi$ and $n(x_0)=N$, so that

$$A=Ne^{-\boldsymbol{e}\phi/\boldsymbol{kT}}.$$

When $i\neq 0$, $V(x_0)=\phi+V$, therefore

$$\begin{aligned} i\int_0^{x_0} e^{\boldsymbol{e}V/\boldsymbol{kT}}\,dx&=\boldsymbol{e}ND e^{-\boldsymbol{e}\phi/\boldsymbol{kT}}(1-e^{-\boldsymbol{e}V/\boldsymbol{kT}})\\ &=\boldsymbol{kT}\sigma e^{-\boldsymbol{e}\phi/\boldsymbol{kT}}(1-e^{-\boldsymbol{e}V/\boldsymbol{kT}})/e. \qquad \ldots\ldots(27)\end{aligned}$$

The integral on the left hand side depends chiefly on the field at the surface and varies much more slowly with V and T than the terms on the right. In fact, for $V\ll\phi$, when the barrier is altered little by the applied voltage, equations (24) and (27) are indistinguishable. With the usual assumption that $\boldsymbol{e}\phi\gg E_0$, V has the form given by equation (23), i.e.

$$\int_0^{x_0} e^{\boldsymbol{e}V/\boldsymbol{kT}}\,dx=\int_0^{x_0} e^{2\pi N\boldsymbol{e}^2(x^2-2xx_0)/\kappa\boldsymbol{kT}}\,dx.$$

The value of the integrand is a maximum at $x=0$, and decreases rapidly as x increases. When $x_0^2>1/4\pi N\boldsymbol{e}^2$, the integral has the approximate form

$$\int_0^{x_0} e^{\boldsymbol{e}V/\boldsymbol{kT}}\,dx=\int_0^{x_0} e^{-4\pi N\boldsymbol{e}^2xx_0/\kappa\boldsymbol{kT}}\,dx.$$

Since practically the whole potential drop takes place in the blocking layer, values of $x\geqslant x_0$ are equivalent to $x=\infty$. The integral can then be evaluated and gives

$$\int_0^{\infty} e^{-4\pi N\boldsymbol{e}^2xx_0/\kappa\boldsymbol{kT}}\,dx=\kappa\boldsymbol{kT}/4\pi N\boldsymbol{e}^2x_0. \qquad \ldots\ldots(28)$$

The current is then

$$i=\sigma\sqrt{\{8\pi N\boldsymbol{e}\kappa(\phi+V)\}}(1-e^{-\boldsymbol{e}V/\boldsymbol{kT}})e^{-\boldsymbol{e}\phi/\boldsymbol{kT}}. \qquad \ldots\ldots(29)$$

Both equations (24) and (27) show rectification in the direction commonly observed.

The contact between a metal and a crystal of the type considered in § 1 must take place through a surface film. Both types of barrier will then be present, and the contact may be represented by figure 10.

The total potential drop across the barrier falls into two parts, V_1 from $x=-\delta$ to $x=0$, and V_2 from $x=0$ to $x=x_0$. Since there is no change on the surface, the field at $x=0$ must be the constant field in the insulator and therefore

$$\left.\begin{aligned} V_1&=F_0\delta_0=4\pi N\boldsymbol{e}x_0/\kappa,\\ V_2&=\int_0^{x_0}F\,dx=2\pi N\boldsymbol{e}^2x_0^2/\kappa,\end{aligned}\right\} \qquad \ldots\ldots(30)$$

where

$$V_1+V_2=\phi+V=\phi_1-\phi_2+V.$$

ϕ_1 and ϕ_2 may be assumed to be so large that transfer of electrons takes place entirely by tunnel effect. For small voltages, i.e. $|V|<\phi$, the current calculated on the diode theory is

$$i=(\sigma/N)\sqrt{(2\pi \boldsymbol{m}/\boldsymbol{kT})}e^{-2\kappa\phi_2^{1/2}\delta}[e^{-e(\phi-V_1)/kT}-e^{-eV_2/kT}]. \quad \ldots\ldots(31)$$

On diffusion theory it is

$$i=\{ss'/(s+s')\}[e^{-e(\phi-V_1)/kT}-e^{-eV_2/kT}]e^{-2\kappa\phi_2^{1/2}\delta}, \quad \ldots\ldots(32)$$

where

$$s=\sigma\sqrt{\{4\pi Ne^2\kappa(\phi+V)\}},$$

$$s'=\sigma/N\sqrt{(2\pi)}.$$

Equations (31) or (32) reduce to (24) or (27) when $V_1=0$, i.e. $\delta=0$, and to (21) when $V_2=0$, i.e. $\phi_1\sim\phi_2$, and, therefore, $x_0=0$. In the intermediate case, rectification may be in either direction, but will not be so pronounced, and the current in the direction of high resistance will not saturate at $V=\boldsymbol{kT}$. This is illustrated by figure 11.

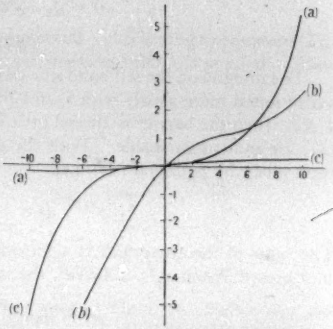


Figure 11. Alternative types of rectifying curves.
(*a*) $i=\text{const.}\ (e^{eV/kT}-1)$.
(*b*) $i=\text{const.}\ \{e^{-e(\phi-V_1)/kT}-e^{-eV_2/kT}\}$ $(\delta/\lambda=1;\ x_0/\lambda=5)$.
(*c*) $i=\text{const.}\ (1-e^{-eV/kT})$.

The equations (21), (31) and (32) are derived on the assumption that an appreciable current is obtained under a small potential difference. If either ϕ or δ is large this is not so. Then a voltage must be applied to the contact, which is comparable to the contact potential difference before an observable current is obtained. The current from a metal, through a surface barrier in strong fields, is given by the Nordheim-Fowler formula (1928)

$$i=\frac{\boldsymbol{e}}{2\pi\boldsymbol{h}}\frac{\mu^{1/2}}{(\phi_1+\mu)\phi^{1/2}}\left(\frac{V}{\delta}\right)^2 e^{-4\kappa\phi_2^{3/2}\delta/3V}. \quad \ldots\ldots(33)$$

From the semiconductor it is

$$i=(\sigma/N)\sqrt{(2\pi \boldsymbol{m}/\boldsymbol{kT})}e^{-4\kappa\phi_2^{3/2}\delta/3V}, \quad \ldots\ldots(34)$$

When $\phi_2<\phi_1$, the latter current will be the greater. Wilson's theory then gives the right direction of rectification.

Since the space-charge is eliminated when $V<-\phi$, the resistance of the contact vanishes in Bethe's theory and there remains only the resistance of the semiconductor. For $V\geqslant \boldsymbol{kT}/\boldsymbol{e}$, equation (31) shows that the current reaches the value

$$i=(\sigma/N)\sqrt{(2\pi \boldsymbol{m}/\boldsymbol{kT})}, \quad \ldots\ldots(35)$$

which is independent of the applied voltage. In strong fields, however, the barrier is distorted by the image force, so that for $V \sim \phi$

$$i = (\sigma/N)\sqrt{(2\pi \boldsymbol{m}/\boldsymbol{k}T)}e^{-e^{3/2}F^{1/2}/\kappa^{1/2}\boldsymbol{k}T}. \qquad \ldots\ldots(36)$$

A voltage such that $V_1 = -\phi$, applied to the surface film contact, removes the space-charge barrier, so that the current is limited only by the barrier of the surface film. For $-\phi_1 \ll V < -\phi$,

$$i = (\sigma/N)\sqrt{(2\pi \boldsymbol{m}/\boldsymbol{k}T)}e^{-2\kappa\phi_2^{1/2}\delta(1-(V-\phi)/4\phi_1)}, \qquad \ldots\ldots(37)$$

and for $V \ll -\phi_1$

$$i = (\sigma/N)\sqrt{(2\pi \boldsymbol{m}/\boldsymbol{k}T)}e^{-4\kappa\phi_2^{1/2}\delta/3(V-\phi)} \qquad \ldots\ldots(38)$$

when $V_1 \geqslant \phi_1$ electrons are drawn from below the Fermi level of the metal. The number of electrons crossing against the field is negligible, and the current is therefore given by

$$i = e\int_{\mu-(\phi+V_1)}^{\mu} N(w)P(w)\,dw, \qquad \ldots\ldots(39)$$

where

$$P(w) \sim e^{-\{2\boldsymbol{h}(c-w)^{1/2}(1-v/4(c-w))\}},$$

$$N(w) \sim (4\pi \boldsymbol{me}/\boldsymbol{h}^3)(\mu - w).$$

Integrating, this gives

$$i = \frac{4\pi \boldsymbol{me}^2}{\boldsymbol{h}^3} \cdot \frac{e^{-\alpha(1-V_1/4\phi_1)}}{\alpha^2(1+V_1/4\phi_1)^2}(2\phi_1)^2\,[1 - e^{-\alpha(1+V_1/4\phi_1)(V_1-\phi)/2\phi_1}\{1+\alpha(1+V_1/4\phi_1)\}], \qquad \ldots\ldots(40)$$

where

$$\alpha = 2\kappa\phi_1^{1/2}\delta.$$

When $V_1 \geqslant \phi_1$, equation (33) may be applied, with V replaced by V_1, i.e.

$$i = \frac{\boldsymbol{e}}{2\pi\boldsymbol{h}} \cdot \frac{\mu^{3/2}}{(\phi_1+\mu)\phi_1^{1/2}}\left(\frac{V_1}{\delta}\right)^2 e^{-4\kappa\phi_2^{3/2}\delta/3V_1}, \qquad \ldots\ldots(41)$$

where, from equation (30),

$$V_1/V = 2x_0\delta/\lambda^2.$$

This set of expressions, together with (31), covers the whole range of voltage for the surface film contact. Figures 12 (*a*), (*b*) and (*c*) have been drawn for suitable values of the constants. The number of unknown parameters, N, ϕ_1, ϕ_2 and δ, is too great to allow a quantitative comparison with experiment, but certain conclusions can be drawn from the shape of the curves obtained.

Figures 13, 14 and 15 are examples of experimental curves obtained for three different kinds of crystals. They are compared with the theoretical curves to which they most closely correspond. If the values assumed for N, ϕ_1 and ϕ_2 were of the right order of magnitude, it would appear that any film (if it exists) on germanium crystals is less than 5 A. thick and has no influence on the rectification. On the silicon crystals it is between 5 and 10 A. and on silicon carbide between 10 and 20 A. thick. The effect of increasing film thickness, for $|V| < \phi$, is to decrease the degree of rectification and to prevent the saturation of current in the direction of high resistance.

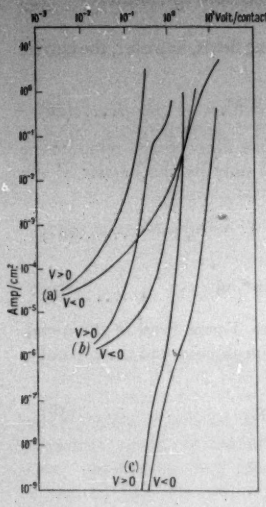


Figure 12. The current-voltage characteristic of the contact between a metal and a semiconductor, for three thicknesses of the surface film.

$\phi_1 = 2{\cdot}0$ volt. (*a*) $\delta/\lambda = 0{\cdot}5$.
$\phi_2 = 1{\cdot}5$ volt. (*b*) $\delta/\lambda = 1{\cdot}0$.
$N = 10^{19}$ cm^{-3}. (*c*) $\delta/\lambda = 2{\cdot}0$.

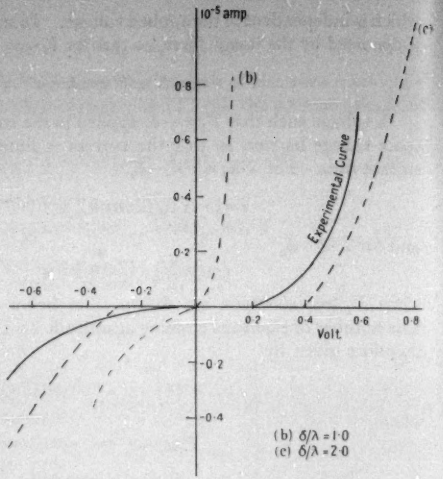


Figure 13. Curve showing rectification in the region $V \sim \phi$ at the contact between silicon carbide and a metal. Experimental data reproduced from a report by E. W. J. Mitchell and R. W. Sillars, by permission of Metropolitan-Vickers Electrical Co. Ltd.

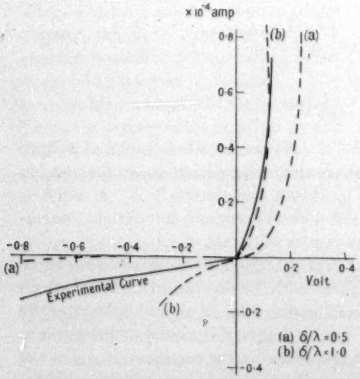


Figure 14. Curve showing rectification in the region V at the contact between silicon and a metal. Experimental data supplied by British Thomson-Houston Co. Ltd.

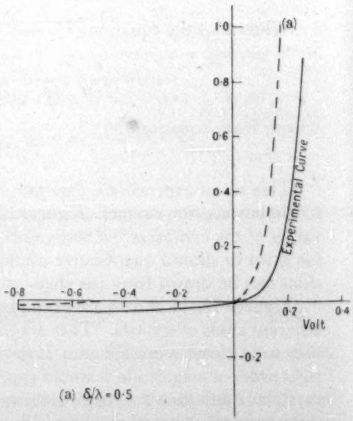


Figure 15. Curve showing rectification in the region V at the contact between germanium and a metal. Experimental data supplied by British Thomson-Houston Co. Ltd.

If the temperature T is changed, the thickness x_0 of the space-charge barrier, and therefore the critical ratio $\delta : x_0$ changes proportionally. Where the thermionic factor $e^{-e\phi/kT}$ also enters, the total temperature effect is complex and is not considered here. For the germanium crystals, where δ may be taken as zero, one has more simply

$$i \propto \sigma e^{-e\phi/kT} = \text{const} \,.\, e^{-E_a/2kT}\, e^{-e\phi/kT},$$

as in Bethe's theory.

§4. CONCLUSION

The theory, applied in §1 to the contact between identical semiconductors, accounts qualitatively for the behaviour of the contacts between a semiconductor and a metal, if the effect of space-charge is taken into account. It follows that both the Wilson and Schottky mechanisms are necessary to a complete treatment of crystal rectifiers.

REFERENCES

BETHE, 1942, Theory of the boundary layer of crystal rectifiers, *B.L. Report*, No. 43–12.
BRAUN and BUSCH, 1942, *Helv. Phys. Acta*, **15**, 571.
FINCH and WILMAN, 1937, *Trans. Faraday Soc.*, **33**, 337.
FOWLER and NORDHEIM, 1928, *Proc. Roy. Soc.* A, **119**, 173.
HEINE and SCHERRER, 1940, *Helv. Phys. Acta*, **13**, 489.
JEANS, *Electricity and Magnetism*, p. 358 (Cambridge University Press).
KURTSCHATOW *et al.*, 1935, *Phys. Z. Sowjet*, **7**, 257.
MITCHELL and SILLARS, 1947, *Metropolitan-Vickers Research Reports* (February).
MOTT, 1939, *Proc. Roy. Soc.* A, **171**, 27.
SCHOTTKY, 1939, *Wiss. Veröff. Siemens-Werke*, **18**, 225.
WILSON, 1932, *Proc. Roy. Soc.* A, **136**, 487.

The Specific Heat of a Linear Ionic Lattice

BY B. DONOVAN *

Imperial College, London

* Now at the University of Sheffield.

Communicated by M. Blackman ; MS. received 12 *June* 1947 ;
in amended form 16 *September* 1947

ABSTRACT. The specific heat at moderately high temperatures has been calculated for a one-dimensional lattice in which both electrostatic and repulsive forces are taken into account. The Thirring expansion method, which does not necessitate a knowledge of the density of normal vibrations of the lattice, has been applied and the resulting values of the specific heat are compared with those obtained by graphical integration of the frequency spectrum.

Thirring's method involves expanding the specific heat in a series in which the terms include summations over the $2n$th powers of the normal frequencies. It is found that values of n greater than 4 lead to computations which are excessively laborious. Nevertheless, with four terms, it is shown that reliable results may be obtained over a considerable range of temperature. At lower temperatures the Thirring method becomes progressively less accurate, showing that the number of terms used is insufficient, and a correction is derived from considerations of the remainder of the series after four terms.

§ 1. INTRODUCTION

THE properties of a one-dimensional lattice were first investigated by Born and von Kármán (1912) who worked out the distribution of normal vibrations on the assumption of quasi-elastic binding forces between neighbouring particles. The specific heat of such a lattice was evaluated by Blackman (1935), who compared his results with those derived from the continuum theory of Debye (1912). The case of the ionic lattice was examined by Broch (1936) who, using the generalized zeta-function, worked out the vibrational spectrum of a one-dimensional lattice taking into account the electrical interactions of the charged particles and the repulsive forces.

In this paper the specific heat of a linear ionic lattice is calculated by two methods. If the vibrational spectrum is represented by a function $\rho(\nu)$ giving the number of frequencies between ν and $\nu+d\nu$, the specific heat at constant volume can be obtained in the form

$$c_v = \boldsymbol{k}\int_0^{\nu_{\max}} \rho(\nu) E\left(\frac{\boldsymbol{h}\nu}{\boldsymbol{k}T}\right) d\nu, \qquad \ldots\ldots(1)$$

where

$$E(x) = \frac{x^2 e^x}{(e^x-1)^2},$$

$\nu_{\max}$ is the maximum frequency, $\boldsymbol{h}$ is Planck's constant, $\boldsymbol{k}$ Boltzmann's constant and T the absolute temperature. Once the function $\rho(\nu)$ is known, the solution of (1) using graphical integration is relatively straightforward, and greater interest attaches to the application of a method due to Thirring (1913) by which the specific heat of moderately high temperatures may be obtained without a knowledge of the vibrational spectrum.

In this method the mean energy of a linear oscillator

$$\bar{u} = \tfrac{1}{2}\boldsymbol{h}\nu + \boldsymbol{k}T\frac{\boldsymbol{h}\nu/\boldsymbol{k}T}{\exp(\boldsymbol{h}\nu/\boldsymbol{k}T)-1},$$

is expanded in powers of $\boldsymbol{h}\nu/\boldsymbol{k}T$, giving

$$\bar{u} = \boldsymbol{k}T\{1 - \sum_{n=1}^{\infty}(-1)^n (B_n/(2n)!)(\boldsymbol{h}\nu/\boldsymbol{k}T)^{2n}\},$$

where B_n are the Bernoulli numbers; the specific heat, obtained by differentiation, is thus expressible as a series of terms in ν^2, ν^4 etc., which have to be summed over all the frequencies of the lattice. These terms become progressively more laborious to compute, and it is desirable to determine whether a comparatively small number of terms will yield useful results. In particular, it is important to know the temperature range over which the method is reliable, and the degree of accuracy to be expected. Some information may thus be obtained for the three-dimensional case.

We consider a chain of particles which may be taken to consist of a number of cells each containing two particles. The coordinate of the kth particle in the lth cell in the equilibrium configuration is given by

$$x_k^l = x_k + la,$$

where x_k is the coordinate of the kth particle in the base cell and a is the lattice constant. The distance between the k'th particle in the base cell and the kth particle in the lth cell is denoted by $x_{kk'}^l$, and the total potential energy of two

particles with coordinates x_k^l and $x_{k'}$ is denoted by $\phi_{kk'}^l$. The equations of motion then assume the form (Born 1923)

$$\left.\begin{aligned}\left\{\omega^2 m_1+\begin{bmatrix}1 & 1\\ x & x\end{bmatrix}\right\}U_1+\begin{bmatrix}1 & 2\\ x & x\end{bmatrix}U_2=0,\\ \begin{bmatrix}2 & 1\\ x & x\end{bmatrix}U_1+\left\{\omega^2 m_2+\begin{bmatrix}2 & 2\\ x & x\end{bmatrix}\right\}U_2=0.\end{aligned}\right\} \quad \ldots\ldots(2)$$

The symbols $\begin{bmatrix}k & k'\\ x & x\end{bmatrix}$ are defined by the equations

$$\begin{bmatrix}k & k'\\ x & x\end{bmatrix}=\sum_{l=-\infty}^{+\infty}(\phi_{kk'}^l)_{xx}\exp\{-2\pi i h x_{kk'}^l/a\}, \quad \ldots\ldots(3)$$

where

$$(\phi_{kk'}^l)_{xx}=\{d^2(\phi_{kk'}^l)/dx^2\}_{|x_{kk'}^l|}. \quad \ldots\ldots(4)$$

Further, h/a is the wave-number, ω is the angular frequency, U_1 and U_2 are the amplitudes and m_1 and m_2 the masses of the two particles.

The potential $\phi_{kk'}^l$ may be conveniently resolved into two parts (Broch 1937), subsequently referred to as B.) : one due to the Coulomb forces and the other due to the repulsive forces. The van der Waals attractive forces are neglected. The functions $\begin{bmatrix}k & k'\\ x & x\end{bmatrix}$ are therefore similarly separable, and for the electrostatic parts one obtains (B., equations (10), (19) and (27)) :

$$\begin{bmatrix}1 & 2\\ x & x\end{bmatrix}^E=\begin{bmatrix}2 & 1\\ x & x\end{bmatrix}^E=-\frac{2\epsilon^2}{a^3}e^{\pi i h}\sum_{l=-\infty}^{+\infty}\frac{\exp(2\pi i h l)}{|l+\frac{1}{2}|^3} \quad \ldots\ldots(5)$$

and

$$\begin{bmatrix}1 & 1\\ x & x\end{bmatrix}^E=\begin{bmatrix}2 & 2\\ x & x\end{bmatrix}^E=\frac{24\epsilon^2}{a^3}\zeta(3)+\frac{2\epsilon^2}{a^3}\sum_{l=-\infty}^{+\infty}{}'\frac{\exp(2\pi i h l)}{|l|^3}, \quad \ldots\ldots(6)$$

where the accent on the summation sign signifies the omission of the term $l=0$. Here $\zeta(t)$ is the Riemann zeta-function and the charge on the jth particle is $+\epsilon$ or $-\epsilon$ according as j is odd or even.

Assuming that the repulsive potential varies as the inverse sth power of the distance, and taking into account nearest neighbours only, the repulsive parts $\begin{bmatrix}k & k'\\ x & x\end{bmatrix}^R$ may be expressed in the form (B., equations (34), (35), (37) and (48))

$$\begin{bmatrix}1 & 2\\ x & x\end{bmatrix}^R=\begin{bmatrix}2 & 1\\ x & x\end{bmatrix}^R=\frac{8\epsilon^2}{a^3}\cos\pi h(s+1)\log 2 \quad \ldots\ldots(7)$$

and

$$\begin{bmatrix}1 & 1\\ x & x\end{bmatrix}^R=\begin{bmatrix}2 & 2\\ x & x\end{bmatrix}^R=-\frac{8\epsilon^2}{a^3}(s+1)\log 2. \quad \ldots\ldots(8)$$

Returning to the equations of motion (2), we see that the normal frequencies are obtained as roots of the determinantal equation

$$\begin{vmatrix}\omega^2 m_1+A(h) & B(h)\\ B(h) & \omega^2 m_2+A(h)\end{vmatrix}=0, \quad \ldots\ldots(9)$$

where

$$A(h)=\frac{8\epsilon^2}{a^3}[3\zeta(3)-(s+1)\log 2]+\frac{2\epsilon^2}{a^3}\sum_{l=-\infty}^{+\infty}{}'\frac{\exp(2\pi i h l)}{|l|^3} \quad \ldots\ldots(10)$$

and

$$B(h)=\frac{8\epsilon^2}{a^3}\cos\pi h(s+1)\log 2-\frac{2\epsilon^2}{a^3}e^{\pi i h}\sum_{l=-\infty}^{+\infty}\frac{\exp(2\pi i h l)}{|l+\frac{1}{2}|^3}. \quad \ldots\ldots(11)$$

In order to consider the distribution of the frequencies we specialize to a chain consisting of N cells, i.e. of $2N$ particles. Application of periodic boundary conditions then gives

$$\exp(2\pi i N h) = 1.$$

Now $0 \leqslant h \leqslant 1$ (since the wavelength cannot be less than the lattice constant); consequently we have

$$h = \mu/N, \qquad \ldots\ldots(12)$$

where μ is any positive integer such that $0 \leqslant \mu \leqslant N$. The allowed values of h have a constant spacing $(=1/N)$ and the normal frequencies are distributed uniformly in h-space, each value of h corresponding to two frequencies.

In the method of Thirring (1913) the heat capacity of a linear oscillator (of frequency ν) is expressed in the form

$$c_\nu = \boldsymbol{k}\left\{1 + \sum_{n=1}^{\infty} (-1)^n \frac{B_n(2n-1)}{(2n)!}\left(\frac{\boldsymbol{h}\nu}{\boldsymbol{k}T}\right)^{2n}\right\},$$

and the series is convergent if $\boldsymbol{h}\nu/\boldsymbol{k}T < 2\pi$. In this expression B_n are the Bernoulli numbers.

In order to derive the heat capacity of the one-dimensional lattice, it is necessary to sum this expression over all the $2N$ frequencies of the lattice, i. e.

$$c_\nu = \boldsymbol{k}\left\{2N + \sum_{n=1}^{\infty} (-1)^n \frac{B_n 2n-1}{(2n)!}\left(\frac{\boldsymbol{h}}{\boldsymbol{k}T}\right)^{2n} \sum_{\nu} \nu^{2n}\right\}. \qquad \ldots\ldots(13)$$

We thus have to perform the summations over all the normal frequencies using equations (9), (10) and (11). For the case in which the two types of particle have equal masses, we write $m_1 = m_2 = m_0$ and the solutions of (9) become

$$\left.\begin{aligned} 4\pi^2 m_0 \nu_1^2 &= -A(h) - B(h), \\ 4\pi^2 m_0 \nu_2^2 &= -A(h) + B(h). \end{aligned}\right\} \qquad \ldots\ldots(14)$$

From these solutions the sums $\sum_{\nu} \nu^{2n}$ may be evaluated, and the cases corresponding to $n = 1, 2, 3, 4$ are treated individually in the next four sections. For $n > 4$ the labour involved is prohibitive and a discussion of the error incurred in neglecting higher order terms is given later. The equations (14) show how two frequencies are associated with every value of h; the frequency curve splits into two branches but these two solutions are identical apart from a difference in phase (B., p. 493).

§ 2. EVALUATION OF $\Sigma \nu^2$

From (14) we obtain

$$\nu_1^2 + \nu_2^2 = -(1/2\pi^2 m_0) A(h),$$

and, summing over both branches to include all frequencies, we can express the required sum as

$$\sum_{\nu} \nu^2 = -(1/2\pi^2 m_0) \sum_{h} A(h), \qquad \ldots\ldots(15)$$

where the right-hand side is to summed over all the permissible values of h. Now from (12) it is clear that the summation over h is equivalent to a summation over μ and, since N is very large, this may be replaced by the integral

$$\int_0^N A(\mu/N)\, \mathrm{d}\mu, \qquad \text{i.e. by} \qquad \int_0^1 N A(h)\, dh.$$

Reference to (10) shows that $A(h)$ consists of two parts, the first independent of h and the second an infinite series involving h. The latter is uniformly convergent in the interval $0 \leqslant h \leqslant 1$ and may therefore be integrated term by term. This, however, gives zero for all integral values of l and we are thus left with

$$\Sigma\nu^2 = -(4N\epsilon^2/\pi^3 m_0 a^3)[3\zeta(3) - (s+1)\log 2]. \qquad \ldots\ldots(16)$$

To obtain the numerical value of the expression in square brackets (subsequently denoted by C) the exponent s of the repulsive forces is given the value 10. Further,

$$\zeta(3) = \sum_{n=1}^{\infty} 1/n^3 = 1{\cdot}20205\ldots,$$

so that

$$C = -4{\cdot}01845. \qquad \ldots\ldots(17)$$

Substituting in (16) gives

$$\Sigma\nu^2 = 1{\cdot}629\, N(\epsilon^2/m_0 a^3). \qquad \ldots\ldots(18)$$

§ 3. EVALUATION OF $\Sigma\nu^4$

From (14) we obtain

$$\sum_{\nu}\nu^4 = N/8\pi^4 m_0^2\left[\int_0^1 A^2(h)\,dh + \int_0^1 B^2(h)\,dh\right]. \qquad \ldots\ldots(19)$$

Now

$$A^2(h) = 64\epsilon^4 a^{-6}C^2 + 32\epsilon^4 a^{-6}C\sigma_1 + 4\epsilon^4 a^{-6}\sigma_1^2, \qquad \ldots\ldots(20)$$

where

$$\sigma_1 = \sum_l{}' \frac{\exp(2\pi i h l)}{|l|^3}.$$

Of these three terms the first is independent of h, the second provides a zero contribution and the third may be expressed as a double sum by means of Cauchy's theorem concerning the multiplication of absolutely convergent series. Upon integration the non-zero terms reduce to

$$\sum_{n=-\infty}^{+\infty}{}' 1/n^6 = 2\sum_{n=1}^{\infty} 1/n^6,$$

and, since

$$\sum_{n=1}^{\infty} 1/n^6 = 1{\cdot}01734\ldots,$$

we have

$$\int_0^1 A^2(h)\,dh = 1{\cdot}042\times 10^3\,\epsilon^4/a^6. \qquad \ldots\ldots(21)$$

Squaring expression (11) gives

$$B^2(h) = 64\epsilon^4 a^{-6}[(s+1)\log 2]^2\cos^2\pi h - 32\epsilon^4 a^{-6}(s+1)\log 2\cos\pi h.\sigma_2 + 4\epsilon^4 a^{-6}.\sigma_2^2, \qquad \ldots\ldots(22)$$

where

$$\sigma_2 = \sum_l \frac{\exp[\pi i h(2l+1)]}{|l+\frac{1}{2}|^3}.$$

Here the first term is straightforward, and the second splits into two sums both of which give, when integrated, a numerical value of 8. Application of Cauchy's theorem to the third term in (22) leads to the expression

$$\sum_{n=-\infty}^{+\infty} \frac{1}{|n+\frac{1}{2}|^6} = 2\sum_{n=0}^{\infty} \frac{1}{(n+\frac{1}{2})^6} = 126\,\zeta(6),$$

using (71) with $x=6$.

Hence $$\int_0^1 B^2(h)\,dh = 421{\cdot}2\,\epsilon^4/a^6, \qquad \ldots\ldots(23)$$

and, inserting (21) and (23) into (19), we have

$$\Sigma\nu^4 = 1{\cdot}877N(\epsilon^2/m_0a^3)^3. \qquad \ldots\ldots(24)$$

§ 4. EVALUATION OF $\Sigma\nu^6$

From (14) we obtain

$$\Sigma_{\nu}\nu^6 = -(N/32\pi^6m_0^3)\left[\int_0^1 A^3(h)\,dh + 3\int_0^1 A(h)B^2(h)\,dh\right]. \qquad \ldots\ldots(25)$$

Now

$$A^3(h) = 512\epsilon^6a^{-9}C^3 + 384\epsilon^6a^{-9}C^2\sigma_1 + 96\epsilon^6a^{-9}C{\sigma_1}^2 + 8\epsilon^6a^{-9}{\sigma_1}^3. \qquad \ldots\ldots(26)$$

Here the first term is constant, the second vanishes when integrated, and the value of the third may be obtained from § 3. The fourth term reduces to the double sum

$$\sum_m{}' \sum_n{}' [|m+n|^3|m|^3|n|^3]^{-1}, \qquad (m+n\neq 0),$$

which is shown in Appendix 1 to have the numerical value 0·816. Integration of (26) thus gives

$$\int_0^1 A^3(h)\,dh = -3{\cdot}400\times 10^4\,\epsilon^6/a^9. \qquad \ldots\ldots(27)$$

For the second term in (25) we find

$$\begin{aligned} A(h)B^2(h) = {} & 512\epsilon^6a^{-9}C[(s+1)\log 2)^2\cos^2\pi h + 32\epsilon^6a^{-9}C{\sigma_2}^2 \\ & - 256\epsilon^6a^{-9}C(s+1)\log 2\cos\pi h.\sigma_2 + 128\epsilon^6a^{-9}[(s+1)\log 2]^2\cos^2\pi h.\sigma_1 \\ & + 8\epsilon^6a^{-9}\sigma_1{\sigma_2}^2 - 64\epsilon^6a^{-9}(s+1)\log 2\cos\pi h.\sigma_1\sigma_2. \qquad \ldots\ldots(28) \end{aligned}$$

Of these six terms the first is straightforward and the second and third may be evaluated using the results of § 3. The fourth term splits into three sums: one gives a zero contribution and the other two have the same value, namely unity. The fifth term may be expanded as a triple sum which reduces to

$$\sum_m \sum_n [|m+n+1|^3|m+\tfrac{1}{2}|^3|n+\tfrac{1}{2}|^3]^{-1}, \qquad (m+n\neq -1),$$

and by adding the fifty largest terms a numerical value of 139 is obtained.

The sixth term in (28) may be written as two double sums which lead to the expression

$$4\left[\sum_m [|m+1|^3|2m+1|^3]^{-1} + \sum_m{}' [|m|^3|2m+1|^3]^{-1}\right].$$

It is easily seen that each of these terms is equivalent to

$$\sum_{m=1}^{\infty} [m^3(2m+1)^3]^{-1} + \sum_{m=1}^{\infty} [m^3(2m-1)^3]^{-1},$$

and the value of this combination (obtained from the general case in Appendix 2) is

$$80 - 48\zeta(2),$$

i.e. $1{\cdot}0432$ $(\zeta(2) = 1{\cdot}6449\ldots.)$.

Integration of (28) thus gives

$$\int_0^1 A(h)B^2(h)\,dh = -1{\cdot}278\times 10^4\,\epsilon^6/a^9, \qquad \ldots\ldots(29)$$

and substituting (27) and (29) into (25) yields

$$\Sigma\nu^6 = 2{\cdot}351\,N(\epsilon^2/m_0a^3)^3. \qquad \ldots\ldots(30)$$

§ 5. EVALUATION OF $\Sigma\nu^8$

From (14) we obtain

$$\Sigma\nu^8 = (N/128\pi^8 m_0^4)\left[\int_0^1 A^4(h)\,dh + 6\int_0^1 A^2(h)B^2(h)\,dh + \int_0^1 B^4(h)\,dh\right]. \quad \ldots\ldots(31)$$

Now

$$A^4(h) = 4096\epsilon^8 a^{-12}C^4 + 4096\epsilon^8 a^{-12}C^3\sigma_1 + 1536\epsilon^8 a^{-12}C^2\sigma_1^2 + 256\epsilon^8 a^{-12}C\sigma_1^3 + 16\epsilon^8 a^{-12}\sigma_1^4. \quad \ldots\ldots(32)$$

Here the first term is constant, the second vanishes when integrated, the third is evaluated in § 3 and the fourth in § 4 (using Appendix 1). The fifth term reduces to the triple sum

$$\sum_l{}' \sum_m{}' \sum_n{}' [|l|^3|m|^3|n|^3|l+m+n|^3]^{-1}, \qquad (l+m+n\neq 0),$$

and addition of the forty largest terms gives a numerical value of 6·79. Integration of (32) thus gives

$$\int_0^1 A^4(h)\,dh = 11{\cdot}178\times 10^5\,\epsilon^8/a^{12}. \quad \ldots\ldots(33)$$

The second term in (31) is obtained from

$$\begin{aligned} A^2(h)B^2(h) = {} & 4096\epsilon^8 a^{-12}C^2[(s+1)\log 2]^2\cos^2\pi h \\ & + 256\epsilon^8 a^{-12}C^2\sigma_2^2 - 2048\epsilon^8 a^{-12}C^2(s+1)\log 2\cos\pi h.\sigma_2 \\ & + 256\epsilon^8 a^{-12}[(s+1)\log 2]^2\cos^2\pi h.\sigma_1^2 + 16\epsilon^8 a^{-12}\sigma_1^2\sigma_2^2 \\ & - 128\epsilon^8 a^{-12}(s+1)\log 2\cos\pi h.\sigma_2\sigma_1^2 \\ & + 2048\epsilon^8 a^{-12}C[(s+1)\log 2]^2\cos^2\pi h.\sigma_1 + 128\epsilon^8 a^{-12}C\sigma_1\sigma_2^2 \\ & - 1024\epsilon^8 a^{-12}C(s+1)\log 2\cos\pi h.\sigma_1\sigma_2. \end{aligned} \quad \ldots\ldots(34)$$

Of these nine terms the first may be integrated straightforwardly and the second and third evaluated using the results of § 3. The fourth term splits into three double sums which lead to the expression

$$\tfrac{1}{4}\left[\sum_m{}' [|m|^3|m+1|^3]^{-1} + \sum_m{}' [|m|^3|m-1|^3]^{-1} + 2\sum_m{}' |m|^{-6}\right].$$

This reduces to $\sum_{m=1}^{\infty} [m^3(m+1)^3]^{-1} + \sum_{m=1}^{\infty} m^{-6}$ (cf. equation (63))

and, using the results of Appendix 1, the numerical values are found to be

$$0{\cdot}1304 + 1{\cdot}0173 = 1{\cdot}1477.$$

The fifth term in (34) may be expanded as a quadruple sum which reduces to

$$\sum_l{}' \sum_m{}' \sum_n [|l|^3|m|^3|n+\tfrac{1}{2}|^3|l+m+n+\tfrac{1}{2}|^3]^{-1},$$

and by adding the forty largest terms a numerical value of 309 is obtained. The sixth term in (34) splits into two triple sums and gives

$$4\left[\sum_m{}' \sum_n{}' [|m|^3|n|^3|2m+2n+1|^3]^{-1} + \sum_m{}' \sum_n{}' [|m|^3|n|^3|2m+2n-1|^3]^{-1}\right],$$

in which both terms have the same value, since changing the signs of m and n in either sum gives the other. By direct addition of the forty largest terms this value is found to be 2·345.

The seventh, eighth and ninth terms in (34) are similar to the fourth, fifth and sixth terms respectively in (28) and hence integration of (34) gives

$$\int_0^1 A^2(h)B^2(h)\,dh = 3{\cdot}900\times 10^5\,\epsilon^8/a^{12}. \qquad \ldots\ldots(35)$$

Lastly we consider the third term in (31), which is obtained from

$$B^4(h) = 4096\epsilon^8a^{-12}[(s+1)\log 2]^4\cos^4\pi h - 4096\epsilon^8a^{-12}[(s+1)\log 2]^3\cos^3\pi h\,.\,\sigma_2 + 1536\epsilon^8a^{-12}[(s+1)\log 2]^2\cos^2\pi h\,.\,\sigma_2^2 - 256\epsilon^8a^{-12}(s+1)\log 2\cos\pi h\,.\,\sigma_2^3 + 16\epsilon^8a^{-12}\sigma_2^4. \qquad \ldots\ldots(36)$$

The first term here is straightforward and the second, written as four separate sums, may be shown to have the value

$$1/8\,[8/27+24+24+8/27] = 6{\cdot}074.$$

The third term may be expressed as three double sums which reduce in the usual way to

$$16\left[\sum_l[|2l+1|^3|2l+3|^3]^{-1} + \sum_l[|2l+1|^3|2l-1|^3]^{-1} + 2\sum_l|2l+1|^{-6}\right]. \qquad \ldots\ldots(37)$$

In this expression the first two terms have the same value (the second emerges from the first on replacing l by $-(l+1)$).

Now $$\sum_l[|2l+1|^3|2l+3|^3]^{-1} = 1+2\sum_{l=1}^{\infty}(2l-1)^{-3}(2l+1)^{-3},$$

and $$\sum_{l=1}^{\infty}(2l-1)^{-3}(2l+1)^{-3} = \sum_{l=1}^{\infty}(4l^2-1)^{-3} = (32-3\pi^2)/64 = 0{\cdot}03736$$

(Knopp 1928, p. 269).

Further, from (71)

$$\sum_l|2l+1|^{-6} = 63\,\zeta(6)/32,$$

so that the numerical value of (37) is 98·483.

The fourth term in (36) splits into two triple sums which lead to

$$256\left[\sum_m\sum_n[|2m+1|^3|2n+1|^3|2m+2n+3|^3]^{-1} + \sum_m\sum_n[|2m+1|^3|2n+1|^3|2m+2n+1|^3]^{-1}\right],$$

and by adding the forty largest terms the value of each sum is found to be 3·12.

To evaluate the fifth term in (36) we have to consider the expression

$$2^{12}\sum_l\sum_m\sum_n[|2l+1|^3|2m+1|^3|2n+1|^3|2l+2m+2n+3|^3]^{-1},$$

and by direct addition, ignoring all terms in which l, m, n are greater than unity, an approximate value of $2{\cdot}86\times 10^4$ is obtained.

Integration of (36) thus gives

$$\int_0^1 B^4(h)\,dh = 2{\cdot}957\times 10^5\,\epsilon^8/a^{12}, \qquad \ldots\ldots(38)$$

and substituting (33), (35) and (38) into (31) finally gives

$$\Sigma\nu^8 = 3{\cdot}091\,N(\epsilon^2/m_0a^3)^4. \qquad \ldots\ldots(39)$$

§6. CALCULATION OF SPECIFIC HEAT BY THIRRING'S METHOD

The specific heat, which may now be computed from (13) by inserting the results given by (18), (24), (30) and (39), is given by

$$c_v = Nk\left\{2-\left(\frac{1\cdot629}{12}\right)\Lambda+\left(\frac{1\cdot877}{240}\right)\Lambda^2-\left(\frac{2\cdot351}{6048}\right)\Lambda^3+\left(\frac{3\cdot091}{172800}\right)\Lambda^4\right\}, \qquad \ldots\ldots(40)$$

where $$\Lambda=(h/kT)^2(\epsilon^2/m_0a^3).$$

Since values of the specific heat obtained from this formula are to be compared with the results of graphical integration it is essential to ensure that the maximum frequency is the same in both cases. The angular frequency may be expressed in terms of the wave number by means of the relation (B, p. 493)

$$\begin{aligned}m_0\omega^2 = {} & 8\epsilon^2a^{-3}(s+1)\log 2(1-\cos\pi h)-24\epsilon^2a^{-3}\zeta(3)\\ & -4\epsilon^2a^{-3}\left\{-\frac{2\pi}{3}+2\sum_{l=1}^{\infty}\left[\frac{\exp(-\pi l^2)}{l^2}+\frac{G(\sqrt{\pi}l)}{2l^3}\right]\cos 2\pi hl\right.\\ & \left.+\pi\sum_{l=-\infty}^{+\infty}[\exp(-\pi[h+l]^2)+\pi[h+l]^2\,\mathrm{Ei}(-\pi[h+l]^2)]\right\}\\ & +4\epsilon^2a^{-3}\left\{2\sum_{l=0}^{\infty}\left[\frac{\exp(-\pi[\tfrac{1}{2}+l]^2)}{(\tfrac{1}{2}+l)^2}+\frac{G(\sqrt{\pi}[\tfrac{1}{2}+l])}{2(\tfrac{1}{2}+l)^3}\right]\cos 2\pi h(\tfrac{1}{2}+l)\right.\\ & \left.+\pi\sum_{l=-\infty}^{+\infty}(-1)^l[\exp(-\pi[h+l]^2)+\pi[h+l]^2\,\mathrm{Ei}(-\pi[h+l]^2)]\right\}. \qquad \ldots\ldots(41)\end{aligned}$$

In this expression $$G(x)=1-F(x),$$

where $F(x)$ is the Gauss error integral, and $\mathrm{Ei}(-x)$ is defined by

$$-\mathrm{Ei}(-x)=\int_x^{\infty}\frac{e^{-t}}{t}\,dt \qquad (0<x<\infty).$$

The ω-h curve (B., p. 498) shows that there is a maximum value for the frequency when h is approximately 0·88 *; the density of normal vibrations (given by $dh/d\omega$) therefore becomes infinite at this point. The behaviour of the density function in the neighbourhood of the maximum frequency, though not troublesome from the physical point of view, provides difficulties in graphical integration, since it is not possible in this region to estimate accurately the area under the density curve.

To avoid this difficulty we approximate to the exact frequency spectrum by replacing the high-frequency end by a single "line" (i.e. by an Einstein term) whose frequency is equal to that corresponding to $h=1$. Figure 1 shows this spectrum together with the exact relation according to equation (41). The range of h over which the Einstein term extends is decided by reference to the density curve, and the upper limit for the graphical integration has been taken to correspond to $h=0{\cdot}675$.

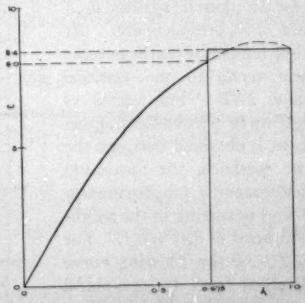


Figure 1. The angular frequency ω as a function of h. The scales are identical with those of B, figure 1.

* The exact position of the maximum depends on the exponent s of the repulsive forces.

The frequency ν_m of the Einstein term and the maximum frequency ν_0 of the continuous spectrum may be determined from (41). The series in this expression are rapidly convergent and numerical calculation is a fairly simple matter. Further, the functions $F(x)$ and $Ei(-x)$ are both available in numerical tables. * The calculations give

$$\nu_m = 1{\cdot}1769(\epsilon^2/(m_0a^3))^{\frac{1}{2}}, \qquad \ldots\ldots(42)$$

$$\nu_0 = 1{\cdot}1208(\epsilon^2/(m_0a^3))^{\frac{1}{2}}. \qquad \ldots\ldots(43)$$

In the Thirring calculation the maximum frequency is adjusted to have the value given by (43). At this stage the parameter θ_0 is introduced, defined by

$$\theta_0 = h\nu_0/k, \qquad \ldots\ldots(44)$$

and the variation of the specific heat with temperature is obtained by plotting c_v as a function of θ_0/T. From (43) and (44) we have

$$\left(\frac{h}{kT}\right)^2\left(\frac{\epsilon^2}{m_0a^3}\right) = \left(\frac{\theta_0}{1{\cdot}121\,T}\right)^2, \qquad \ldots\ldots(45)$$

and this expression, substituted in (40), enables the specific heat to be calculated for any value of θ_0/T.

Table 1 shows the values obtained by this method over a range of temperature extending down to $\theta_0/T = 4{\cdot}0$. The table also shows the results of the graphical method discussed in § 8. The results are shown graphically in figures 2 and 3, in which the specific heat is plotted against θ_0/T and T/θ_0 respectively. At high temperatures the specific heat tends to the classical value $2Nk$. For values of θ_0/T up to 3·0 excellent agreement is obtained between the two methods, the maximum discrepancy (approximately 0·5%) occurring in the neighbourhood of $\theta_0/T = 0{\cdot}7$. For $\theta_0/T > 3{\cdot}0$ the Thirring curve diverges from the graphical curve and gives higher values for the specific heat. This indicates that the formula (40)

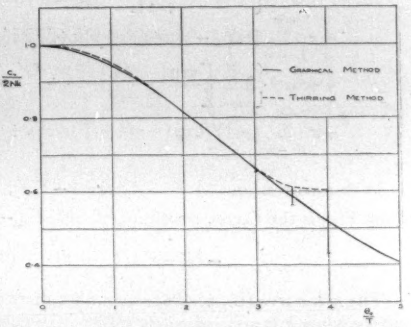


Figure 2. The specific heat as a function for θ_0/T. The corrections given in § 7 are shown for $\theta_0/T = 3{\cdot}0$, 3·5 and 4·0.

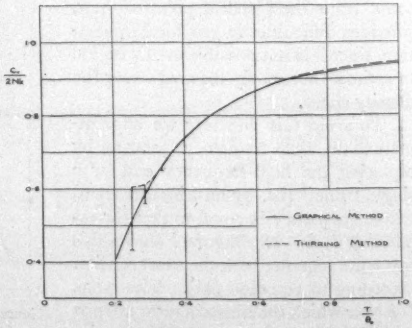


Figure 3. The specific heat as a function of T/θ_0. The corrections given in § 7 are shown for $\theta_0/T = 3{\cdot}0$, 3·5 and 4·0.

* e.g. Publications of the Federal Works Agency, Works Projects Administration (New York 1940–41).

is valid up to $\theta_0/T = 3{\cdot}0$; above this value the results become rapidly less accurate, as might be expected, since the infinite series in (13) has been cut off at the fourth term.

Table 1

θ_0/T	0	0·3	0·7	1·0	1·25	1·5	2·0	2·5	3·0	3·5	4·0	5·0
$c_v/2N\boldsymbol{k}$ (Thirring)	1·000	0·995	0·974	0·948	0·921	0·890	0·818	0·741	0·667	0·611	0·604	
$c_v/2N\boldsymbol{k}$ (graphical)	1·000	0·993	0·969	0·945	0·918	0·892	0·817	0·740	0·667	0·587	0·519	0·407

§ 7. CONSIDERATION OF REMAINDER IN THIRRING'S METHOD

Having found the range of θ_0/T over which equation (40) can yield useful results, it is of interest to investigate the remainder after four terms in the original expression (13). In this section a rough estimate is made of the upper limit to the remainder in terms of θ_0/T.

An extension of Euler's summation formula is used, starting with the identity (Knopp 1928, p. 528)

$$\frac{\alpha}{e^\alpha-1}+\frac{\alpha}{2}\equiv 1-\sum_{n=1}^{\lambda}(-1)^n\frac{B_n}{(2n)!}\alpha^{2n}+\frac{\alpha^{2\lambda+2}}{e^\alpha-1}\int_0^1 P_{2\lambda+1}(x)e^{\alpha x}\,dx, \qquad \ldots\ldots(46)$$

where

$$P_{2\lambda+1}(x)=(-1)^{\lambda-1}\sum_{m=1}^{\infty}(2\sin 2m\pi x)/(2m\pi)^{2\lambda+1}. \qquad \ldots\ldots(47)$$

Now the mean energy of a linear oscillator is given by

$$\bar{u}=\tfrac{1}{2}\boldsymbol{h}\nu+\boldsymbol{h}\nu/[\exp(\boldsymbol{h}\nu/\boldsymbol{k}T)-1],$$

and hence the indentity (46), with $\alpha=\boldsymbol{h}\nu/\boldsymbol{k}T$, shows that the remainder after λ terms is

$$\boldsymbol{k}T\left\{\alpha^{2\lambda+2}(e^\alpha-1)^{-1}\int_0^1 P_{2\lambda+1}(x)e^{\alpha x}\,dx\right\}.$$

Differentiating this expression we obtain the remainder $R_\lambda =$ in the expansion for the specific heat:

$$R_\lambda=\boldsymbol{k}\alpha^{2\lambda+2}(e^\alpha-1)^{-1}\int_0^1 P_{2\lambda+1}(x)e^{\alpha x}\,dx$$

$$-(\boldsymbol{h}\nu/T)[(2\lambda+2)(e^\alpha-1)\alpha^{2\lambda+1}-e^\alpha\alpha^{2\lambda+2}](e^\alpha-1)^{-2}\int_0^1 P_{2\lambda+1}(x)e^{\alpha x}\,dx$$

$$-(\boldsymbol{h}\nu/T)\alpha^{2\lambda+2}(e^\alpha-1)^{-1}\int_0^1 xP_{2\lambda+1}(x)e^{\alpha x}\,dx. \qquad \ldots\ldots(48)$$

To determine the maximum value of $|R_\lambda|$, with $\lambda=4$, we consider the first two terms together and write them as

$$\boldsymbol{k}\{\alpha^{10}(e^\alpha-1)^{-1}-(10\alpha^{10}(e^\alpha-1)-\alpha^{11}e^\alpha)(e^\alpha-1)^{-2}\}\int_0^1 P_9(x)e^{\alpha x}\,dx. \qquad \ldots\ldots(49)$$

An upper limit to the value of the integral may be derived from the inequality

$$|P_9(x)|\leqslant(2/(2\pi)^9)\sum_{m=1}^{\infty}m^{-9},$$

which follows from (47). Since

$$\sum_{m=1}^{\infty} m^{-9} < \sum_{m=1}^{\infty} m^{-8}, \quad \text{and} \quad \sum_{m=1}^{\infty} m^{-8} = \pi^8/9450,$$

$|P_9(x)|$ is certainly less than the quantity $\hat{P}$, where

$$\hat{P} = (256 + 9450\pi)^{-1}. \qquad \ldots\ldots(50)$$

Thus for the limiting value of $\int_0^1 P_9(x)e^{\alpha x}\,dx$ we may write

$$\hat{P}\int_0^1 e^{\alpha x}\,dx = \hat{P}\alpha^{-1}(e^{\alpha} - 1),$$

which, substituted in (49), gives

$$\boldsymbol{k}\hat{P}\{-9\alpha^9 + \alpha^{10} + \alpha^{10}/(e^{\alpha} - 1)\}. \qquad \ldots\ldots(51)$$

Since the last term contains $e^{\alpha} - 1$ in the denominator, we may write $e^{\alpha} - 1 = \alpha$ to obtain the form

$$\boldsymbol{k}\hat{P}\alpha^9(\alpha - 8),$$

which indicates that (51) is negative, since the convergence of formula (13) requires $\alpha < 6$.

We now examine the third term in (48), viz.

$$\boldsymbol{k}\left\{\alpha^{11}(e^{\alpha} - 1)^{-1}\int_0^1 xP_9(x)e^{\alpha x}\,dx\right\}, \qquad \ldots\ldots(52)$$

and seek to establish the sign of the integral since this term has to be subtracted from the first two. The function $P_9(x)$ has the form sketched in figure 4 (Knopp 1928, p. 537).

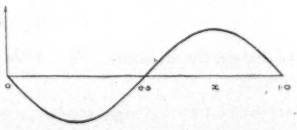


Figure 4. Form of the function $P_9(x)$.

In view of the fact that the integral contains the factor $e^{\alpha x}$, where α is positive, it is clear that the integral itself, and hence the expression (52), must be positive.

To obtain the maximum value of $|R|_{\lambda=4}$, therefore, the maximum value of (52) must be subtracted from (51). The limiting value of the integral in (52) may be expressed as

$$\hat{P}\int_0^1 xe^{\alpha x}\,dx = \hat{P}\alpha^{-2}\{e^{\alpha}(\alpha - 1) + 1\},$$

which, substituted in (52), gives

$$\boldsymbol{k}\hat{P}\{\alpha^{10}e^{\alpha}(e^{\alpha} - 1)^{-1} - \alpha^9\}.$$

Subtracting this expression from (51) we obtain $-8\boldsymbol{k}\hat{P}\alpha^9$ so that we may write

$$|R| < 8\boldsymbol{k}\hat{P}(\boldsymbol{h}\nu/\boldsymbol{k}T)^9,$$

which, summed over all possible frequencies, becomes

$$|R| < 8\boldsymbol{k}\hat{P}(\boldsymbol{h}/\boldsymbol{k}T)^9\Sigma\nu^9. \qquad \ldots\ldots(53)$$

Now $\Sigma\nu^9 < \nu_0 \Sigma\nu^8$ and hence, using (39) and (43), we can define a quantity $|R|_{\max.}$ by the equation

$$|R|_{\max.} = 8N(1{\cdot}12 \times 3{\cdot}09)\boldsymbol{k}\hat{P}(\boldsymbol{h}/\boldsymbol{k}T)^9(\epsilon^2/m_0a^3)^{9/2}.$$

Finally we substitute from (45) and (50) and obtain

$$(|R|_{\max.}/2N\boldsymbol{k}) = (1{\cdot}82 \times 10^{-6})[\theta_0/1{\cdot}12T]^9. \qquad \ldots\ldots(54)$$

Since R has been shown to be negative, this equation gives an upper limit to the correction which must be subtracted from the values given in table 1.

Values of the remainder calculated from (54) are given in table 2 and the corrections are indicated by vertical lines in figures 2 and 3.

The correction increases rapidly with decreasing temperature, since it is proportional to the ninth power of θ_0/T, and although negligible for $\theta_0/T < 2{\cdot}0$ it is nearly 2% at $\theta_0/T = 3{\cdot}0$. This is not borne out by the results obtained (table 1) and it is clear from the somewhat drastic approximations that the upper limit to $|R|$ given by (54) is considerably in excess of the true correction. For the two cases $\theta_0/T = 3{\cdot}5$ and $4{\cdot}0$ the values given in table 2 are approximately twice the discrepancies between the "Thirring" and "Graphical" curves in figures 2 and 3.

Table 2

θ_0/T	1·0	2·0	3·0	3·5	4·0
$\lvert R\rvert_{\max.}/2Nk$	5×10^{-7}	3×10^{-4}	$0{\cdot}01_3$	$0{\cdot}05_1$	$0{\cdot}17_1$

§8. CALCULATION OF SPECIFIC HEAT BY GRAPHICAL INTEGRATION

The fundamental expression (1) may be evaluated by graphical means if the density of normal vibrations is known. In the present case, with the frequency spectrum illustrated in figure 1, the specific heat may be expressed in the form

$$c_v = \boldsymbol{k}\int_0^{\nu_0} \rho(\nu)\mathrm{E}\left(\frac{\boldsymbol{h}\nu}{\boldsymbol{k}T}\right)d\nu + 2N\boldsymbol{k}t_{\mathrm{E}}\mathrm{E}\left(\frac{\boldsymbol{h}\nu_m}{\boldsymbol{k}T}\right), \qquad \ldots\ldots(55)$$

where the function $\rho(\nu)$ is normalized by means of the relation

$$\int_0^{\nu_0} \rho(\nu)\,d\nu = 2Nt_{\mathrm{c}}. \qquad \ldots\ldots(56)$$

Here t_{E} and t_{c} are weight factors which represent the fractions of the total number of vibrations included in the Einstein and continuous regions respectively. Since equal ranges of h contain equal numbers of vibrations, we have

$$t_{\mathrm{c}} = 0{\cdot}675, \qquad t_{\mathrm{E}} = 0{\cdot}325. \qquad \ldots\ldots(57)$$

The frequency spectrum for the region $0 \leqslant \nu \leqslant \nu_0$ has been calculated from equation (41) and the values of ω and h so obtained have been used to calculate the density function $\rho(\nu)$ in terms of ω by means of the relation (B, p. 499)

$$\begin{aligned}\frac{\omega}{\rho} = \frac{4\pi\epsilon^2}{a^3m_0}\Bigg\{&(s+1)\log 2\sin\pi h + 2\sum_{l=1}^{\infty}\left[\frac{\exp(-\pi l^2)}{l} + \frac{\mathrm{G}(\sqrt{\pi l})}{2l^2}\right]\sin 2\pi hl \\ &-\pi\sum_{l=-\infty}^{+\infty}[1-(-1)^l](h+l)\,\mathrm{Ei}(-\pi[h+l]^2) \\ &-2\sum_{l=0}^{\infty}\left[\frac{\exp(-\pi[\tfrac{1}{2}+l]^2)}{\tfrac{1}{2}+l} + \frac{\mathrm{G}(\sqrt{\pi[\tfrac{1}{2}+l]})}{2(\tfrac{1}{2}+l)^2}\right]\sin 2\pi h(\tfrac{1}{2}+l)\Bigg\}. \qquad \ldots\ldots(58)\end{aligned}$$

For any value of θ_0/T we find can $\mathrm{E}(h\nu/kT)$ as a function of ω; computation is facilitated by means of the relation

$$\mathrm{E}(x) = x^2e^x/(e^x-1)^2 = \tfrac{1}{2}x^2/(\cosh x - 1). \qquad \ldots\ldots(59)$$

Hence the product $\rho(\nu)\mathrm{E}(h\nu/kT)$ can be obtained as a function of ω. In this way curves have been constructed for values of θ_0/T varying from 0·3 to 5·0, and the areas under the curves determined by counting squares. The second term in (55) may be calculated immediately from (59), since

$$\theta_m/T = 1{\cdot}05\,\theta_0/T,$$

and the sum of the two contributions gives the specific heat, as listed in table 1.

Referring to figures 2 and 3, the slight discrepancy in the neighbourhood of $\theta_0/T = 0{\cdot}7$ is rather puzzling, but could probably be decreased by raising the frequency ν_m of the Einstein term or increasing its weight t_{E}. At the low temperature end of the curves one would expect the graphical method to give reasonably accurate results, since in this region the specific heat is controlled by the low-frequency end of the spectrum and hence the effect of the approximation (figure 1) is least noticeable.

ACKNOWLEDGMENT

The author wishes to thank Dr. M. Blackman for suggesting this problem and for his continued interest in it, and for many helpful discussions during the course of the work.

APPENDIX 1

In this appendix a method is given by which an approximate numerical value may be obtained for the expression

$$S_1 = \sum_m{}' \sum_n{}' [|m+n|^3 |m|^3 |n|^3]^{-1} \qquad \ldots\ldots(60)$$

$$(m \neq 0, \quad n \neq 0, \quad m \neq -n).$$

Considering first positive values of n only, and putting $n = \alpha$, we obtain series of the form

$$\alpha^{-3} \sum_m{}' [|m|^3 |m+\alpha|^3]^{-1},$$

and in the two regions $m > 0$ and $m < -\alpha$, we have

$$\alpha^{-3} \sum_{m=1}^{\infty} m^{-3}(m+\alpha)^{-3}, \qquad \ldots\ldots(61)$$

while between the two prohibited values of m (0 and $-\alpha$) there are $(\alpha - 1)$ terms, whose sum may be represented by

$$\Sigma\, \xi^{-3}\eta^{-3}, \qquad \ldots\ldots(62)$$

taken over all non-zero values of ξ and η satisfying the condition $\xi + \eta = \alpha$.

The same reasoning will apply to negative values of n. For if we write $n = -\alpha$ and consider the series

$$\alpha^{-3} \sum_m{}' [|m|^3 |m-\alpha|^3]^{-1},$$

it is clear that the two regions $m > \alpha$ and $m < 0$ will each yield (61) and the remaining $(\alpha - 1)$ terms will be given, as before, by (62).

We thus have the result

$$\sum_m{}' [|m|^3|m+\alpha|^3]^{-1} + \sum_m{}' [|m|^3|m-\alpha|^3]^{-1} = 4 \sum_{m=1}^{\infty} m^{-3}(m+\alpha)^{-3} + 2\Sigma\xi^{-3}\eta^{-3}, \quad \ldots\ldots(63)$$

which we can apply to (60) and obtain

$$S_1 = 4 \sum_{m=1}^{\infty} [m^{-3}(m+1)^{-3}] + \frac{1}{8}\left[2 + 4 \sum_{m=1}^{\infty} m^{-3}(m+2)^{-3}\right] + \frac{1}{27}\left[\frac{1}{2} + 4 \sum_{m=1}^{\infty} [m^{-3}(m+3)^{-3}]\right] + \frac{1}{64}\left[\frac{4}{27} + \frac{1}{32} + 4 \sum_{m=1}^{\infty} m^{-3}(m+4)^{-3}\right] + \ldots. \text{ etc.} \quad \ldots\ldots(64)$$

By expressing $1/[m^3(m+\alpha)^3]$ in partial fractions, we can derive the result

$$\sum_{m=1}^{\infty} \frac{1}{m^3(m+\alpha)^3} = \frac{6}{\alpha^5} \sum_1^{\infty} \left[\frac{1}{m} - \frac{1}{m+\alpha}\right] - \frac{3}{\alpha^4} \sum_1^{\infty} \left[\frac{1}{m^2} + \frac{1}{(m+\alpha)^2}\right] + \frac{1}{\alpha^3} \sum_1^{\infty} \left[\frac{1}{m^3} - \frac{1}{(m+\alpha)^3}\right], \quad \ldots\ldots(65)$$

in which the three terms on the right may all be calculated easily for small values of α. The original expression (60) converges sufficiently rapidly for the sequence (64) to be cut off at, say, $\alpha = 4$.

The sum in the first term in (65) is simply

$$1 + \tfrac{1}{2} + \ldots. + 1/\alpha,$$

while that in the second term is equal to the sum of the first α terms subtracted from $2\sum_1^{\infty} 1/m^2$. In the third sum the terms cancel out with the exception of the first α terms which do not overlap. Values of the expression in (65) for $\alpha = 1, 2, 3, 4$ are given in table 3.

Table 3

α	1	2	3	4
$\sum_{m=1}^{\infty} m^{-3}(m+\alpha)^{-3}$	0·13040	0·03940	0·01687	0·00864

Substituting these values in (64) we obtain

$$S_1 = 0{\cdot}816. \quad \ldots\ldots(66)$$

Referring to (26) it will be seen that $8S_1$ is small compared to $512C^3$ and therefore, in this instance, a more refined calculation of S_1 is unnecessary.

APPENDIX 2

In this appendix a closed expression is developed for the sum of the series

$$S_2(\beta) = \sum_{m=1}^{\infty} m^{-3}(2m+\beta)^{-3}, \quad \ldots\ldots(67)$$

$$(\beta = \pm 1, \pm 3, \ldots.).$$

Using the method of partial fractions we find

$$\sum_{m=1}^{\infty}\frac{1}{m^3(2m+\beta)^3}=\frac{24}{\beta^5}\sum_{1}^{\infty}\left[\frac{1}{m}-\frac{1}{m+\beta/2}\right]-\frac{6}{\beta^4}\sum_{1}^{\infty}\left[\frac{1}{m^2}+\frac{1}{(m+\beta/2)^2}\right]+\frac{1}{\beta^3}\sum_{1}^{\infty}\left[\frac{1}{m^3}-\frac{1}{(m+\beta/2)^3}\right]. \quad \ldots\ldots(68)$$

The three terms on the right will be considered in turn and it will be convenient to deal with positive and negative values of β separately. For the first term, with the notation

$$s(\beta)=\sum_{m=1}^{\infty}\left[\frac{1}{m}-\frac{1}{m+\beta/2}\right],$$

we have

$$s(\beta)-s(1)=\operatorname*{Lim}_{M\to\infty}\left\{\sum_{1}^{M}\left[\frac{1}{m}-\frac{1}{m+\beta/2}\right]-\sum_{1}^{M}\left[\frac{1}{m}-\frac{1}{m+\tfrac{1}{2}}\right]\right\}$$

$$=\operatorname*{Lim}_{M\to\infty}\left\{\sum_{1}^{M}\frac{1}{m+\tfrac{1}{2}}-\sum_{\frac{1}{2}(\beta+1)}^{M+\frac{1}{2}(\beta-1)}\frac{1}{n+\tfrac{1}{2}}\right\},$$

where $n=m+\frac{1}{2}(\beta-1)$.

Thus, for $\beta\geqslant 1$,

$$s(\beta)-s(1)=\sum_{1}^{\frac{1}{2}(\beta-1)}\frac{1}{m+\tfrac{1}{2}}=2\left(\frac{1}{3}+\frac{1}{5}+\ldots.+\frac{1}{\beta}\right).$$

Now $$s(1)=2\left(\frac{1}{2}-\frac{1}{3}+\frac{1}{4}-\ldots.\right)=2(1-\log 2),$$

and hence $$s(\beta)=2\left(\frac{1}{1}+\frac{1}{3}+\ldots.+\frac{1}{\beta}-\log 2\right). \quad \ldots\ldots(69)$$

For $\beta\leqslant -1$, we have

$$s(\beta)-s(1)=-\sum_{\frac{1}{2}(\beta+1)}^{0}\frac{1}{n+\tfrac{1}{2}}=-2\left(\frac{1}{1}-\frac{1}{1}-\frac{1}{3}-\ldots.+\frac{1}{\beta+2}\right),$$

and hence $$s(\beta)=2(1-\log 2)-2\left(\frac{1}{1}-\frac{1}{1}-\frac{1}{3}-\ldots.\frac{1}{\beta+2}\right). \quad \ldots\ldots(70)$$

The second term in (68) may be split up into two parts, of which the first is simply

$$\sum_{m=1}^{\infty}1/m^2=\zeta(2).$$

For the second part we use the relation

$$(1-2^{-x})\zeta(x)=1^{-x}+3^{-x}+5^{-x}+\ldots.(x>1), \quad \ldots\ldots(71)$$

which, for $x=2$, gives

$$\sum_{1}^{\infty}(m+\beta/2)^{-2}=3\zeta(2)-\sum_{1}^{\frac{1}{2}(\beta+1)}(m-\tfrac{1}{2})^{-2}. \quad \ldots\ldots(72)$$

This holds for $\beta\geqslant 1$, while for $\beta\leqslant -1$ the corresponding expression is

$$3\zeta(2)+\sum_{\frac{1}{2}(\beta+1)}^{-1}(m+\tfrac{1}{2})^{-2}. \quad \ldots\ldots(73)$$

The third term in (68) is similar and we have, for the first part,

$$\sum_{m=1}^{\infty}1/m^3=\zeta(3),$$

while for the second part, using (71) with $x=3$, we find

$$\sum_{1}^{\infty}(m+\beta/2)^{-3}\begin{cases}=7\zeta(3)-\sum_{1}^{\frac{1}{2}(\beta+1)}(m-\frac{1}{2})^{-3}, & (\beta\geqslant 1), \quad \ldots\ldots(74)\\ =7\zeta(3)+\sum_{\frac{1}{2}(\beta+1)}^{-1}(m+\frac{1}{2})^{-3}, & (\beta\leqslant -1). \quad \ldots\ldots(75)\end{cases}$$

From (68) the original sum for $\beta\geqslant 1$ is obtained from (69), (72) and (74):

$$S_2(\beta)=\frac{48}{\beta^5}\left(\frac{1}{1}+\frac{1}{3}+\frac{1}{5}+\ldots.+\frac{1}{\beta}-\log 2\right)-\frac{6}{\beta^4}\left(4\zeta(2)-\sum_{1}^{\frac{1}{2}(\beta+1)}\frac{1}{(m-\frac{1}{2})^2}\right)$$
$$+\frac{1}{\beta^3}\left(\sum_{1}^{\frac{1}{2}(\beta+1)}\frac{1}{(m-\frac{1}{2})^3}-6\zeta(3)\right). \quad \ldots\ldots(76)$$

For $\beta\leqslant -1$ the corresponding expression is obtained from (70), (73), and (75):

$$S_2(\beta)=\frac{48}{\beta^5}\left[(1-\log 2)-\left(\frac{1}{1}-\frac{1}{1}-\frac{1}{3}-\ldots.+\frac{1}{\beta+2}\right)\right]$$
$$-\frac{6}{\beta^4}\left(4\zeta(2)+\sum_{\frac{1}{2}(\beta+1)}^{-1}\frac{1}{(m+\frac{1}{2})^2}\right)$$
$$-\frac{1}{\beta^3}\left(\sum_{\frac{1}{2}(\beta+1)}^{-1}\frac{1}{(m+\frac{1}{2})^3}+6\zeta(3)\right). \quad \ldots\ldots(77)$$

For the special case $\beta=1$, (76) reduces to

$$S_2(+1)=80-6[8\log 2+4\zeta(2)+\zeta(3)],$$

and for $\beta=-1$, (77) gives

$$S_2(-1)=6[8\log 2-4\zeta(2)+\zeta(3)].$$

REFERENCES

BLACKMAN, M., 1935, *Proc. Roy. Soc.* A, **148**, 365.
BORN, M. and VON KÁRMÁN, TH., 1912, *Phys. Z.*, **13**, 297.
BORN, M., 1923, *Atomtheorie des festen Zustandes* (Leipzig: Teubner).
BROCH, E. K., 1937, *Proc. Camb. Phil. Soc.*, **33**, 485.
DEBYE, P., 1912, *Ann. Phys., Lpz.*, **39**, 789.
KNOPP, K., 1928, *Theory and Application of Infinite Series* (London: Blackie).
THIRRING, H., 1913, *Phys. Z.*, **14**, 867.

The Interpretation and Application of Electron-Diffraction 'Kikuchi-Line' Patterns—Part I. The Determination of the Crystal Unit Cell, its Orientation and the Crystal Symmetry

BY H. WILMAN

Applied Physical Chemistry Laboratories, Imperial College, London

MS. received 22 *July* 1947

ABSTRACT. The interpretation of electron-diffraction Kikuchi-line patterns from single crystals is developed as a powerful independent and general means of determining the crystal lattice type, dimensions and orientation and the Laue symmetry of the atomic arrangement. This aim is realized by a construction of the reciprocal lattice by methods which are described and illustrated. The satisfactory accuracy of the method is due to the relatively

precise definitions for the location of the Bragg reflection positions, derived and stated here for the first time, which enable the method to be used even when the crystal lattice constants are large so that the bands are narrow and diffuse-edged, as with rhombic sulphur. The patterns shown also illustrate features of the line intensities, especially the reduction in intensity when the bands pass near the undeflected-beam spot.

§ 1. INTRODUCTION

SINGLE-CRYSTAL diffraction patterns provide much more complete though usually also more complex information than is yielded by the ring patterns obtained from polycrystalline material, with respect to the structure of the material, i.e. lattice type, symmetry and atomic arrangement. It is therefore especially important to be able to interpret and utilize single-crystal patterns in such a way as to reveal all the information available in a form easy of application to all cases. The interpretation of the single-crystal Kikuchi-line pattern is developed below as a powerful independent technique by which the structure and orientation of single crystals can be determined.

Electron-diffraction patterns of black and white lines from a single crystal of mica were first observed in 1928 by Kikuchi, who showed the relationship between the line positions and the crystal net planes, and the rough agreement between the relative line intensities and the relative X-ray reflecting powers of the orders of reflection from a series of net planes. Since then, although such patterns have been used as a means of defining the orientation of crystals of known structure (e.g. Thirsk and Whitmore 1940) and of checking the structures of graphite and cadmium iodide (Finch and Wilman 1936, 1937 a) they have otherwise only been made use of in a general way, by inferring from their presence that there are relatively large thicknesses (several hundred A. or more) of coherently diffracting crystals in the specimen, and from their diffuseness or displacement the degree of fine-scale surface roughness, mosaic structure and distortion of the crystal.

In order to use these patterns to determine the lattice structure, orientation and Laue symmetry of a crystal, all that is required is to construct the reciprocal lattice, and this can be achieved in the way described below. Moreover, it can be done independently of whether or not a clear pattern of Laue spots is also present in the pattern, and independently of any other source of knowledge. The crystal lattice and its orientation can be determined from a single photograph if this records a wide enough range of diffracted rays, i.e. the usual 10 to 20° from the primary beam direction.

The method is also of general applicability, for it has been demonstrated (Finch, Quarrell and Wilman 1935, Finch and Wilman 1936, 1937 a, 1937 b) that clear Kikuchi-line patterns can be obtained by transmission through thin cleavage flakes or natural lamellar crystals in favourable cases, or by reflection from (*a*) smooth natural faces, (*b*) cleavage faces, (*c*) 'parting' fractures, (*d*) conchoidal fractures, (*e*) polished crystal surfaces (plane or curved), or (*f*) polished and lightly etched surfaces. It is the only stationary-crystal electron-diffraction method applicable if the patterns contain few or no Bragg or Laue spots owing to the narrowness of the Laue zones. In such cases the crystal surface can usually be slightly roughened by careful abrasion or etching (see examples of FeS_2 and ZnS respectively: Finch and Wilman 1937 b) so as to yield more extensive spot patterns, or the crystal can be rotated during exposure of the plate, but the accuracy

of location of the lattice rows from the spacings and curvatures of the spot rows, i.e. the Laue zones, is in any case often very much less than in the Kikuchi-line method, in which angles between net-planes are directly measurable.

§2. THE BASIS OF THE INTERPRETATION OF KIKUCHI-LINE PATTERNS

(i) *The origin and general nature of the lines and bands*

For the present purpose it suffices to apply, as Kikuchi did, Rutherford and Andrade's (1914) basic conception of the line pairs as originating by selective Bragg reflection by the crystal net-planes from the diffuse background scattering, which has a radial fall of intensity with increasing angle of deviation from the primary beam. A practically parallel pair of black and white lines in the pattern is thus the intersection made with the photographic plate by the two halves of a cone round the net-plane normal, with semi-apical angle $90°-\theta$; θ is the Bragg angle, related to the net-plane spacing d and the electron wavelength λ by Bragg's law, $2d \sin\theta = n\lambda$, where n is an integer, the order number of the diffraction. It is found (Kikuchi 1928, Shinohara 1932, Kirchner 1932, Emslie 1934, Finch, Quarrell and Wilman 1935, Finch and Wilman 1936, 1937 a, 1939) that the measured line-pair separations correspond well (to within about 1 to 2%) with this simplified interpretation, the wavelength of the electrons which contribute to the lines being practically identical with that of the primary beam.

In cases so far examined (Kikuchi 1928, Kirchner 1932) the relative intensities of the different line-pair diffractions also show a general correspondence with the relative structure factors and electron or X-ray reflecting powers of the net-planes when allowance is made for the different inclinations of the planes to the beam. The relative intensities of orders of reflection from a net-plane are, however, appreciably modified by secondary scattering, or more correctly by dynamic interaction of the scattered wavelets, so as to even them out partially. In certain definable cases orders of reflection may appear which are forbidden according to the theory of single-scattering, but this cannot affect extinctions characteristic of the lattice type (Brandenberger 1937).

Kirchner (1932) has pointed out that the elementary diffraction theory for single scattering also leads to a Kikuchi-line breadth proportional to the corresponding plane spacing d/n, and inversely proportional to the thickness of the diffracting part of the crystal in a direction normal to the net-plane. The lines corresponding to d/n about 2 A. or more are in fact generally very broad and diffuse, while those of higher orders from the same plane or from planes with lower d/n are progressively much sharper. The dynamical theory of scattering has been found by Lamla (1938) to give an expression for the line breadth proportional to $1/\sin 2\theta$, which is nearly the same as Kirchner's result, since 2θ is small for fast electrons. Though the theory for selective single scattering from the diffuse background goes so far to account for the position and form of the lines it fails to explain the existence of the prominent bands with symmetrical intensity distribution, which still appear, except near the primary beam, even when the net-plane is parallel to the primary beam. Hayasi (1934), however, showed that such bands are to be expected if from each scattering centre spherical waves spread out with amplitude proportional to the cosine of the angle of deviation from the primary beam, and these are scattered again once by all the other atoms in the crystal. The more complete dynamical theory, applied by Shinohara (1932 a), also yielded an

intensity distribution across the symmetrical bands which was similar to that observed (figure 1 (*b*)). For present purposes, however, these details of the intensity distribution need not be considered further.

(ii) *Location of the Bragg reflection positions and measurement of their separation D*

For present purposes it is first desirable to define more accurately than has hitherto been done the positions of the Bragg reflection directions relative to the Kikuchi lines and band edges, which are very diffuse when they arise from planes whose spacings are large.

In the case of the symmetrical bands associated with planes that are parallel to the primary beam it appears to have been generally assumed that the Bragg positions are at the band edges; these, as the author now defines them, are the points where the intensity is zero relative to the general background in the region considered, as it passes from a positive value (excess over the background) near the maximum on the side within the band, to a negative value which reaches a minimum of intensity slightly farther out. For example, Thomson and Cochrane (1939, pp. 110–116) seem to indicate this to be their opinion, and in indexed patterns of Emslie (1934), Tillman (1935), Finch, Quarrell and Wilman (1935) and Finch and Wilman (1936, 1937 a), the lines denoting the consecutive orders of Bragg reflection from such planes are drawn in accordance with this view. No close quantitative test yet appears to have been made, and this is now supplied by measurements on patterns taken at not too small a camera-length (~22 cm.) and obtained with a finely-focused narrow electron beam (~55 kv.) reflected from a perfect magnesium-oxide cleavage face (cube face (001)), so that the lines are as sharp as possible.

Figure 1 (*a*) shows part of such a pattern where the median of the strong vertical band system due to the (100) plane lies in the plane of incidence, and the undeflected-beam spot lies on this median 3·8 cm. below the lower edge of the figure. The intensity distribution across this band system in a region where obliquely crossing bands are not prominent, is of the form shown in figure 1 (*b*). In this figure are also shown measurements made on the plate at the 400, 600 and 800 band edges, corresponding to points where the intensity was estimated visually to be the same as that of the general background in the same neighbourhood.

It will be seen that there is very good agreement in the spacings for the strong and relatively sharply-defined 400 and 600 band edges, and only slightly less so for the faint but sharp 800. The positions for the very diffuse 200 band edges calculated from these measurements were found also to agree excellently with the points where the intensity was the same as that of the background. These regularly spaced positions are thus evidently the successive orders of the Bragg reflection positions. It is also clearly shown in figure 1 (*a*) that the vertically elongated but laterally very sharp Laue spots lie in every case within 0·001 cm. of these positions when the net-plane passes exactly through the primary beam. (Actually, in figure 1 (*a*) the Kikuchi line pattern is not exactly symmetrical about the plane of incidence, but is displaced sideways by 0·007 cm. relative to the Laue-spot rows, due to an azimuth deviation of 0° 1′ from the symmetrical one, but the quantitative correspondence is clear.) This fact confirms that these are the Bragg reflection positions for electrons which have practically the same speed and energy as those of the incident beam; in the case of the very sharp higher-order band edges such as 800 and 10, 0, 0 the electrons

contributing to the Kikuchi line pattern are calculated to have lost not more than 50 electron-volts energy.

In the case of net-planes not parallel to the primary beam, black and white line-pairs arise in general, but here also there is a band of increased intensity

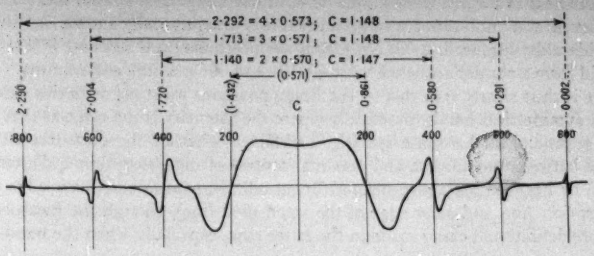


Figure 1 (*b*).

between the lines, which is strong when the line-pair separation is small, corresponding to planes of high spacing. This can be seen prominently in figure 2 (*a*), which shows also the sector-like effect caused by the crossing and superposition of parts of two or more such bands. In these unsymmetrical bands and line-pairs it seems hitherto to have been usual to measure the distance between the centres of the black and white lines, assuming these to occur at the Bragg reflection positions.

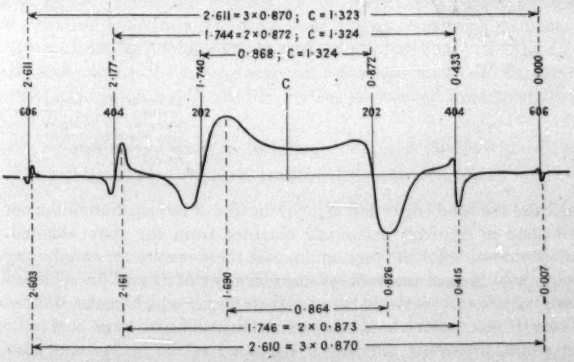


Figure 2 (*b*).

To test this, measurements have now been made on several successive orders of line-pairs or band edges such as those in figure 2 (*a*), which is also part of a pattern from an MgO cleavage face at an azimuth near that of figure 1 (*a*); thus the pattern adjoins on the right that recorded in figure 1 (*a*). An example of the measurements, for the orders of diffraction from the (101) plane, is given in figure 2 (*b*), at the points indicated relative to the intensity distribution shown qualitatively in the figure. It will be seen from the upper series of measurements that the points

where the intensity is the same as the background intensity again form a regularly-spaced series as in the symmetrical bands. On the other hand, the lower series of measurements shows that while the distance apart of corresponding main maxima and minima are virtually identical with corresponding spacings of the upper series, nevertheless as a whole these points do not form a regularly spaced set. The 202 maximum and minimum are so broad that it is unequivocally shown that they are considerably displaced (0·028 cm.) from the positions (at 0·826 and 1·718) which would form a regular sequence with the 404 and 606 maxima and minima.

It is thus clearly seen that (i) the Bragg positions must occur in this case also, as in symmetrical bands, practically where the intensity is the same as that of the background; (ii) the same spacing is obtained whether the measurements are made between the minima and maxima (centres of the absorption and intensified lines) or between the above-defined Bragg reflection positions (*inner* edge of the absorption line and *outer* edge of the intensified line), though the measurement is more definite and easily made in the latter case, especially when the band-width

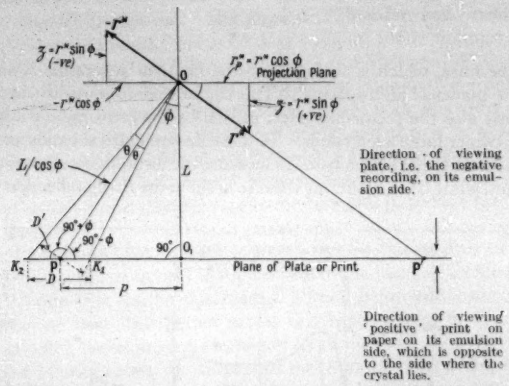


Figure 3.

D is small and the band edges diffuse; (iii) the line of intersection of the net-plane with the plate is relatively accurately obtained from the above-defined Bragg reflection positions. For the present purpose these results are mainly important in showing how greater accuracy of measurement of D can be achieved, thus permitting satisfactory analysis even of patterns in which many diffuse-edged bands occur, from crystals whose lattice axial dimensions are large. Measurement between points where the intensity is the same as the background also often allows relatively accurate measurements to be made, even when some of the lines are much broadened due to a curvature, mosaic structure or distortion of the crystal.

§ 3. CONSTRUCTION OF THE PROJECTION OF THE RECIPROCAL LATTICE FROM THE LINE POSITIONS

(i) *Calculation of reciprocal-lattice point positions from the line positions*

In figure 3 let the crystal origin be at O, near the centre of the diffracting part of the crystal, and let OO_1, of length L, be normal to the plate and meet it at O_1. Let PO_1P' and OP represent in section the plate or the 'positive' print, and a

plane through O parallel to a net-plane system (*hkl*) in the crystal, inclined at ϕ to OO_1. Then the radius vector $\mathbf{r}^*$ from O to the reciprocal-lattice point $[[hkl]]^*$ is $h\mathbf{a}^*+k\mathbf{b}^*+l\mathbf{c}^*$ and is along the normal OP′ to this plane, with length $1/d_{hkl}$. To define the position of this point $[[hkl]]^*$ and the symmetrical point $[[-h,-k,-l]]^*$ in space, it is only necessary to specify the direction and the magnitude $r^*\cos\phi$ of the projection of $\mathbf{r}^*$ on the plane through O normal to OO_1 (parallel to the plate), together with the distance $r^*\sin\phi$ of the point from this plane, which is considered positive if measured from the plane along OO_1 and negative if along O_1O.

These magnitudes can be obtained easily from the camera-length L and measurements of the perpendicular p from O_1 to the line-pair median, and of the separation D of the line-pair K_1K_2, measured along the perpendicular O_1P. Thus

$$r^*=1/(d/n)=\sin\theta/\lambda\simeq\theta/\lambda\simeq(D'\cos\phi)/\lambda L\simeq(D\cos^2\phi)/\lambda L.$$

$$\cos\phi=L/(L^2+p^2)^{\frac{1}{2}};\qquad \sin\phi=p/(L^2+p^2)^{\frac{1}{2}};\qquad \tan\phi=p/L.$$

Whence

$$\begin{aligned}r_p{}^*&=r^*\cos\phi\simeq(D\cos^3\phi)/\lambda L\\&=\left(\frac{D}{\lambda L}\right)\cdot\left\{1-\frac{3}{2}\left(\frac{p}{L}\right)^2+\frac{15}{8}\left(\frac{p}{L}\right)^4-\frac{35}{16}\left(\frac{p}{L}\right)^6+\cdots\right.\\&=(1/\lambda L)\,.\,D\,.\,f(p/L),\\z&=r^*\sin\phi=(p/L)\,.\,r^*\cos\phi.\end{aligned}$$

Figure 4(*a*) can be used to find $f(p/L)$ for any values of p and L likely to be encountered normally, since under the horizontal p/L scale are marked scales of p for a range of L values. The projection of the reciprocal lattice is drawn by plotting as radius vectors $\lambda L\,.\,r^*\cos\phi$, i.e. $D\,.\,f(p/L)$, to a convenient scale, for example a five (or sometimes ten) times scale, the corresponding distances $z\,.\,\lambda L$ of the reciprocal-lattice points from the projection plane $(p/L)\,.\,D\,.\,f(p/L)$, being written by each point as the actual distance *in centimetres* relative to this same scale.

An exact expression for D, i.e. PK_1+PK_2, is obtained from triangles OPK_1 and OPK_2 in figure 5:

$$PK_2=L'\sin\theta/\cos(\phi+\theta);\qquad PK_1=L'\sin\theta/\cos(\phi-\theta),$$

thus

$$\begin{aligned}D=PK_1+PK_2&=L'\sin2\theta\cos\phi/(\cos^2\phi-\sin^2\theta)\\&\simeq L'\sin2\theta/\cos\phi\simeq2L'\sin\theta/\cos\phi\\&=2L\sin\theta/\cos^2\phi.\end{aligned}$$

The last approximate expression for D is accurate to within 0·5% even when ϕ is as much as 20° ($p\sim7{\cdot}7$ cm. when $L=22$ cm.) and θ as much as 4° ($D\sim3$ cm. with $L=22$ cm.), thus the same accuracy, sufficient for present purposes, is obtained in $r^*\cos\phi$, because from this D value we have without further approximation:

$$r^*\cos\phi=(n/d)\cos\phi=(2\sin\theta\cos\phi)/\lambda=(D\cos^2\phi)/\lambda L'=(D\cos^3\phi)/\lambda L,$$

which is the expression already given above.

Similarly, to find how far the trace of the net-plane on the plate is displaced from the actual line-pair 'median', midway between K_1 and K_2 and normal to O_1P, we have

$$\begin{aligned}(PK_2-PK_1)/D&=L'\sin\theta\,.\,[\{1/\cos(\phi+\theta)\}-\{1/\cos(\phi-\theta)\}]/D\\&=2L'\sin^2\theta\,.\,\sin\phi/(\cos^2\phi-\sin^2\theta)D\\&=\tan\theta\tan\phi\\&=(p/L)\tan\theta\simeq D_p/2L(L(+p^2)^{\frac{1}{2}},\end{aligned}$$

by using the above accurate expression for $PK_1 + PK_2 = D$. Since θ is never likely to exceed 6° in normal practical cases, it follows that even if p/L is as high as 0·3 and $\theta = 6°$, the plane trace is displaced from the 'median' by only $(0{\cdot}3 \times 0{\cdot}1)D/2$ i.e. $0{\cdot}03(D/2)$, which can be neglected for present purposes.

A further consideration sometimes required is the variation of line-pair separation D_y with distance y along the 'median' or net-plane trace, since the distance apart D, i.e. K_1K_2, at the hyperbolae vertices ($y = 0$) is sometimes not directly measurable because this region lies outside the recorded photograph. It can be shown that

$$D_y^2 = D^2 + 4y^2/(\cos^2\phi \cot^2\theta - \sin^2\phi),$$

whence

$$(D/D_y)^2 = 1 - (4y^2/D_y^2)\sin^2\theta/(\cos^2\phi - \sin^2\theta).$$

As shown above,

$$D \simeq 2L \sin\theta/\cos^2\phi,$$

thus, neglecting $\sin^2\theta$ in comparison with $\cos^2\phi$ we have, when $D_y \simeq D$,

$$(D/D_y)^2 \simeq 1 - (y^2/L^2)\cos^2\phi = 1 - (y^2/L^2)\{L^2/(L^2 + p^2)\}.$$

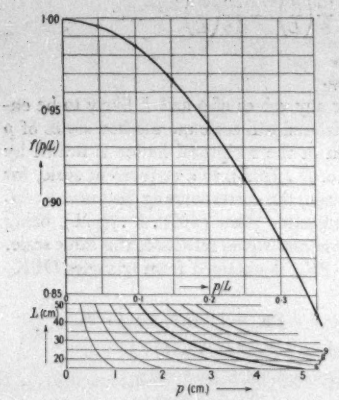


Figure 4 (*a*).

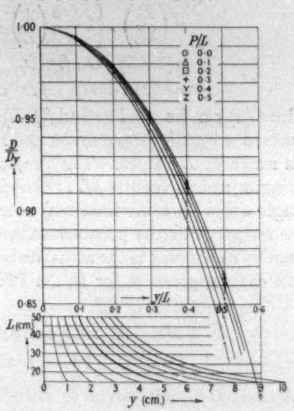


Figure 4 (*b*).

Figure 4 (*b*) shows the values of D/D_y corresponding to this expression, over the range of y/L and p/L concerned, from which can be interpolated in any usual case the factor D/D_y by which D_y must be multiplied to estimate D.

(ii) *The relation between a 'positive' print and the initial 'negative' recording*

It is convenient to measure D and p of the line-pairs on either a 'positive' or a 'negative' reproduction of the initial photographic 'negative' recording of the electron diffraction pattern.

If a negative reproduction is made by first printing a positive on to a plate and taking a print from this on paper, it can be considered in its geometry as identical with the initial negative recording.

A positive reproduction, however, is usually preferred owing to its greater clearness of detail, and it bears in its geometry the relation of a mirror image to the

initial negative. More exactly, if the plate and print are viewed face side up, with top edges roughly in line, and then by rotation of one a line BA in the print is placed in line with the corresponding line AB in the plate, then the print is the mirror image of the plate (apart from interchange of black and white) in a plane which is normal to the line chosen, and cuts it midway between the pair of 'A' or 'B' points.

Alternatively, if the positive print and the plate are put in register with the sensitive plane faces together and are then opened apart by rotating the print about *any* line in their plane until the sensitive faces are again coplanar, then the print and plate are mirror images of each other relative to that line. It follows that, in respect of its geometry the 'positive' print can be regarded as identical with the initial 'negative' which would have been obtained if a crystal had been used which was the mirror image of the crystal actually used, in respect of its lattice, orientation and atomic structure. If the crystal lattice and the atomic structure both possess a plane of symmetry (or have a Laue symmetry containing one) then the details of the positive print are congruent with those of an initial 'negative' recording obtainable from an *actual* crystal if it were in an orientation which is the mirror image of the actual orientation in a plane perpendicular to that of the negative and its positive print (both with sensitive sides up), and passing through the line about which the print was rotated to being it facing the same way up as the negative.

It is this equivalence which has been invoked hitherto in most cases where positive prints of photographs from a crystal known to have a plane of symmetry have been indexed, both Kikuchi-line and spot patterns having been indexed as if they were the initial 'negative' recording (apart from interchange of black and white) viewed along the beam direction from behind the crystal (e.g. in Finch, Quarrell and Wilman 1935, and Finch and Wilman 1935, 1937a, 1937b).

However, it is preferable in such cases, and it is necessary in the case of crystals of unknown structure or those known to have no plane of symmetry, to adopt the following unambiguous system which will be applied in the examples given below. The pattern, whether a positive or a negative, can be thought of as situated so that the features of its intensity distribution coincide with the corresponding distribution of the electrons in the recording plane; a 'negative' print is then to be considered as if it were the actual negative recording viewed along the primary beam direction, from behind the crystal (figure 3), and a 'positive' print is to be regarded (apart from interchange of black and white) as if it were the initial negative recording viewed from behind, in the direction opposite to that of the primary beam, i.e. looking towards the crystal.

(iii) *Practical details of pattern measurement and projection construction*

Location of O_1. Though it is not absolutely essential to locate O_1, the foot of the perpendicular from the crystal to the plate, to within more than about 2 cm. when L is about 20 to 25 cm., it is obvious that the more accurately this is done the more consistent will be the estimated lattice-point positions and the accuracy of determination of the lattice symmetry and dimensions. One indirect way of locating O_1 sufficiently closely is from observations of the positions of the focused undeflected-beam spot F (figure 5), recorded in the photograph, and of the unfocused-beam spot U, the camera-length L, and the distances BA and BC of the focusing-coil from the anode diaphragm A and the plate. Thus, approximately,

by assuming that any change in beam direction caused by the focusing-coil occurs near the central plane of the coil, we have

$$\begin{aligned} CO_1 &\simeq BE + E_1O_1 \\ &\simeq BE + (CF - BE).\{1 - (OO_1/BC)\} \\ &= BE.OO_1/BC + CF\{1 - (OO_1/BC)\} \\ &= (OO_1/BC)(AB/AC).CU + \{1 - (OO_1/BC)\}.CF. \end{aligned}$$

AC, AB, BC, CU and CF can all be measured, and the vectorial form of the expression applies to both the cases (*a*) and (*b*) of figure 5, though that of (*a*) gives the best beam focusing, making F as near U as possible, and it should therefore always be used. If F and U coincide,

$$CO_1 = CF.(AC - OO_1)/AC.$$

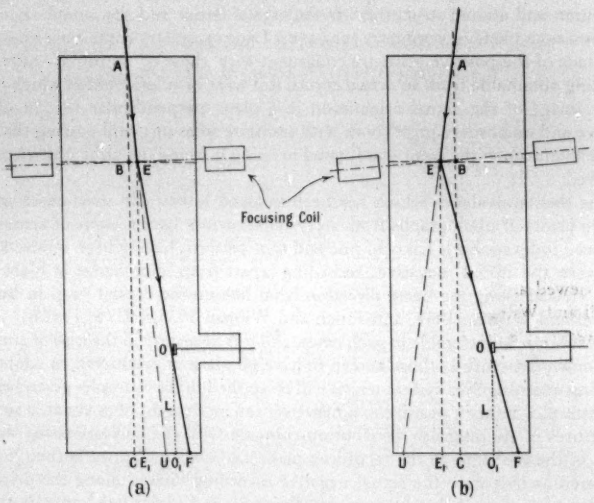


Figure 5.

Measurement of D and p. When a print of the negative has been prepared and stuck down on a suitable-sized piece of paper, the line pairs are sorted out and extensions drawn round the print so as to label each of the line-pair medians with a serial number, as in figures 7(*a*) and 8(*a*). Corresponding black and white lines are practically parallel and of similar intensity and sharpness. The fainter higher orders are not always visible continuously across the whole photograph, but usually appear more clearly on regions of raised intensity due to other band systems. The line-pair separation D at or near the foot of the perpendicular p from O_1 to the lines are then measured on the print, using a thin-edged finely-graduated millimetre scale to avoid parallax, and measuring D usually to an accuracy of 0·1 mm. with the aid of a hand lens magnifying about five times. The perpendicular distances p of the medians from O_1 are also

measured, usually to 0·5 or 0·3 mm., or if possible, to 0·1 mm. in cases where p is only a few millimetres; and a table of D and p values is constructed from which $D \cdot f(p/L)$ and $(p/L)D \cdot f(p/L)$ values can be calculated.

Where possible D should be derived from measurements between the sharper higher orders, divided by the order number. If it is necessary to measure between faint lines, the average of several estimates should be taken. Note especially that when the main line pairs have been sorted out the sharpness of any intensified line can be used to give an indirect estimate of D and so usually leads to identification of the practically parallel absorption line (with the help of a lens), even when this is too faint to be noticed immediately.

If D cannot be measured directly because the region near the foot of the perpendicular from O_1 to the median lies outside the recorded photograph, the spacing D_y must be measured where possible at a distance y from this region, and D estimated by use of figure 4(*b*), as described in § 3 (i).

Construction of the reciprocal-lattice projection. It is desirable to mark points on both sides of the origin, which is chosen some 20 cm. or more to one side of the print, and the geometrical positions of several higher orders should be marked even if they are too faint to be clearly distinguished in the pattern. The serial number of the net-plane is marked below each order, with the order number as suffix if desired. Accuracy of location of the radius-vector normals to the line-pair medians is essential and is assisted by careful use of transparent 60° set-squares of 10-inch hypotenuse. Allowance has to be made for the appreciable opposite curvature of the components of a line-pair in estimating the direction of the median, by finding the direction midway between chords or tangents in the same region of the line-pair.

The final stage is to write above each point in the projection the calculated distance of the point from the projection plane, with the appropriate sign. The projection is to be considered always as viewed from below the recording plane, i.e. in a direction opposite to the primary beam, so that the positive z distances (along OO_1, as defined in § 3 (i)) are from the projection plane towards the observer, and negative distances are on the opposite side. With the sign convention stated above and if a positive print is used, the projection points whose radius vectors are in the same direction as that in which p is measured from O_1 will have negative distances from the projection plane (cf. figure 3).

For the purpose of determining the lattice shape, dimensions and orientation, it is desirable to mark the reciprocal-lattice points as small points surrounded by a ring, but it will often be useful to make copies showing the points as black discs with diameters a measure of the intensities of the diffractions.

(iv) *Use of scales to read off* $\lambda L \cdot r_p^*$ *and* λLz *directly from line positions*

When the line-pair separations D are mostly between 1 and 2 cm., and the patterns not too intricately criss-crossed, time and trouble can be saved with but little loss in accuracy by the use of measuring scales from which $D \cdot f(p/L)$, i.e. $\lambda L \cdot r_p^*$ and $D(p/L) \cdot f(p/L)$, i.e. λLz can be read off directly as the difference in the scale readings at the black-and-white-line Bragg positions when the scale is placed normal to the lines with its zero at O_1.

The possibility of constructing such scales is due to the fact that the factor $f(p/L)$, by which D is multiplied to obtain $\lambda L \cdot r_p^*$, is not far from unity in all cases

likely to be needed, and that its variation is small over a range of p/L of approximately 0·1 corresponding to band widths D up to 2 cm. when $L \simeq 20$ cm. Under these conditions if a function, y, of p is derived such that

$$dy/dp = f(p/L),$$

we can take as a close approximation

$$\Delta y = \Delta p \,.\, f(p/L), \qquad \text{or} \qquad (y_2 - y_1) = (p_2 - p_1) \,.\, f\{(p_1 + p_2)/2L\}$$

even if $p_2 - p_1$ is made so large as to take the value of the line-pair separation D, so that

$$y_2 - y_1 = D \,.\, f\{(p_1 + p_2)/2L\}.$$

Correspondingly, to construct the scale to read off $\lambda L \,.\, z$ directly, we require a function y' given by

$$dy'/dp = (p/L) \,.\, f(p/L).$$

We have thus:

$$f(p/L) = \cos^3\phi = [1 + (p^2/L^2)]^{-3/2},$$

$$y = \int f(p/L)\, dp = p \,.\, f_2(p/L) = p - p\{1 - f_2(p/L)\},$$

$$y' = \int (p/L) \,.\, f(p/L)\, dp = L\{1 - f_2(p/L)\}.$$

The values of $f_2(p/L)$ calculated are given in table 1, whence values of $(p - y)$ and y' were calculated corresponding to a range of p values from $L = 15$, 20, 25, 30, 35, 40, 45 and 50 cm. From these data scales were drawn which were then photographically reduced to the required size. A typical scale is shown in figure 6. Similar scales for any required intermediate value of L can be constructed if necessary in the same way.

Table 1. Variation of $1 - f_2(p/L)$ with p/L

p/L	·05	·1	·15	·2	·25	·3	·35	·4	·45	·5
$1 - f_2(p/L)$	·00125	·0050	·01106	·0194	·0299	·0422	·0561$_5$	·0715	·0880	·1054

A trial of such scales over a wide range of p and D showed excellent agreement with the calculated values of $D \,.\, f(p/L)$ and $(p/L)D \,.\, f(p/L)$. The best scale for practical use is a reproduction on a photographic plate, but it is also possible to use a paper print of the appropriate scale, stuck on the underside of a piece of transparent plastic sheet, if the shrinkage of the paper after processing is found to be negligible. The zero is placed at O_1 then the scale moved slightly if necessary so that a scale mark comes opposite one Bragg position, whence only a single interpolation to a tenth of a division is necessary at the other Bragg position and the difference of the two readings is easily found and is then written down in the list of pattern data. When the list is completed the projection can then be constructed as described above.

§4. CHOICE OF AXES, DETERMINATION OF SYMMETRY AND ORIENTATION, INDEXING OF PATTERN, AND CALCULATION OF LATTICE CONSTANTS

When an adequate number of points are located in the projection and the distances of the lattice points from the projection plane are marked, it is easy to recognize the main reciprocal-lattice rows, because any straight line in space is also represented by a straight line in the projection and, further, the distance $\lambda L z$ of the lattice points from the projection plane must be proportional to their distance along

the line in the projection. The rows with densest lattice-point population are of course the most prominent. Similarly, it is easy to recognize which are the main lattice planes because they contain two or more systems of parallel densely-populated lattice rows. As a convenient 'scaffolding' or reference of coordinates to assist in this process, it is helpful to draw lightly the main lattice lines through the coplanar set of points (forming a two-dimensional lattice) due to the planes of a main zone, usually one whose zone axis is near the primary beam direction, so that the points in the reciprocal lattice plane correspond to the Laue spot pattern in the circular spot zones.

Having tentatively drawn in pencil some coplanar sets of main lattice rows in a favourable part of the projection, so as to show part of the cross-grating set of

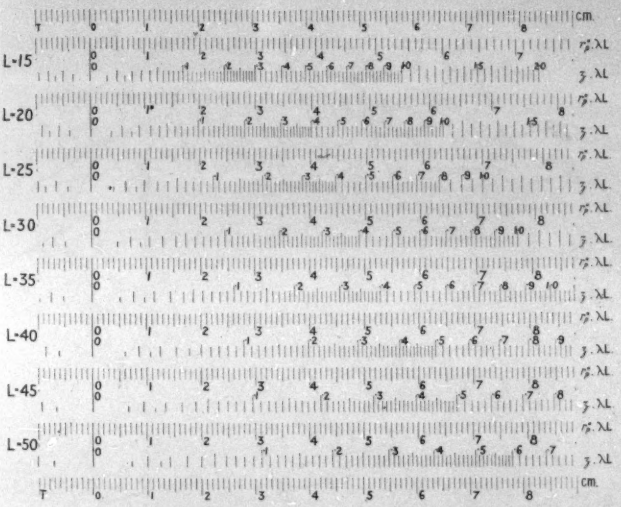


Figure 6.

lattice points in such planes, a different coplanar set can be explored and delineated in the same or another part of the diagram. From such initial trials it is usually clear whether angles of 90° or 120° occur, and which is the appropriate crystallographic system of axes and lattice symmetry, and the whole diagram is then mapped out in lattice rows and planes accordingly, so as to outline the neighbouring unit cells of the reciprocal lattice; the indexing of the lattice points then follows easily as the simple process of writing down the coordinates of the points relative to the chosen axes.

From a consideration of the indices of the observed diffractions, supplemented by similar data from photographs obtained at other crystal settings, any systematic extinction of definite types of diffractions can be recognized and used in the usual way to determine the possible space groups characteristic of the atomic arrangement in the crystal. Allowance must be made, as in x-ray diffraction, for the fact that the "Laue symmetry" shown by diffraction patterns includes a centre of

symmetry even if there is none in the symmetry of the crystal structure; and other data such as observation of crystal forms and etch figures is required to decide this point. In electron diffraction, as already noted, there is also the possibility in certain cases that orders of reflection from a net-plane may occur with appreciable intensity, due to secondary or dynamic scattering, which would otherwise be forbidden according to the more approximate single-scattering theory. The further consideration of Kikuchi-line intensities will be left to a later contribution.

The angles between the reciprocal-lattice rows in planes inclined to the main projection plane, and the spacings between the points in the rows, can be estimated by the usual methods of projective geometry. Side projections in suitable directions, on planes perpendicular to the initial projection plane are especially useful

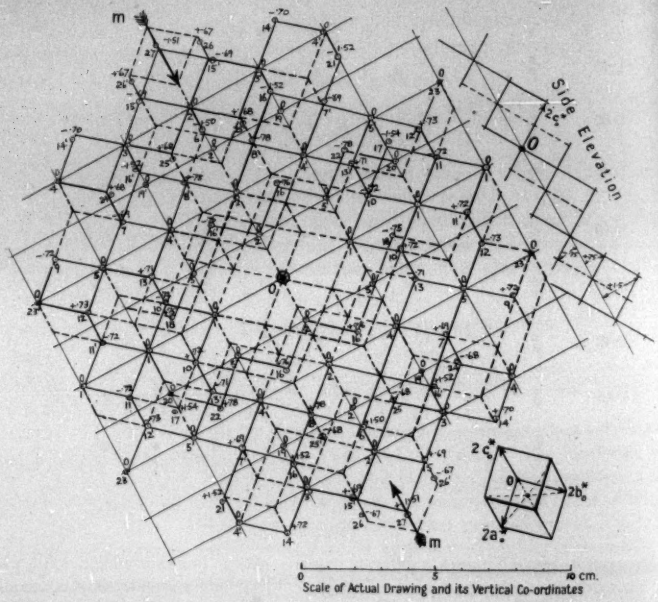


Figure 7 (*b*).

when these can be drawn so as to give desired angles and spacings directly (see figures 7 (*b*) and 8 (*b*)). In awkward cases it may sometimes be most direct to construct the lattice-point positions in space by mounting the projection on cardboard and sticking pins through the projected points so that their heads are at the appropriate heights. For convenience all the points can be brought on to one side of the cardboard base by adding a constant height to the calculated distances of the points from the projection plane.

The reciprocal-lattice angles and the lengths proportional to its axes are estimated from the projection or the model by taking averaged directions and lengths most consistent with the constructed lattice point positions over a region comprising as many unit cells as possible; the actual reciprocal axes are then

Figure 1 (*a*). MgO.

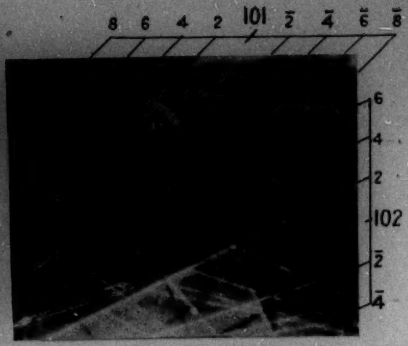


Figure 2 (*a*). MgO.

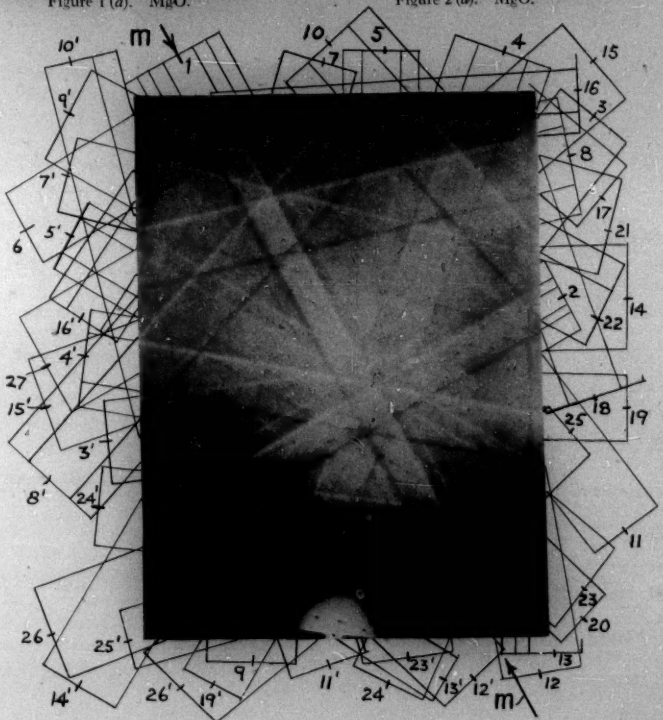


Figure 7 (*a*). NaCl, polished and etched face; L~22 cm., 50 kv.

To face page 354

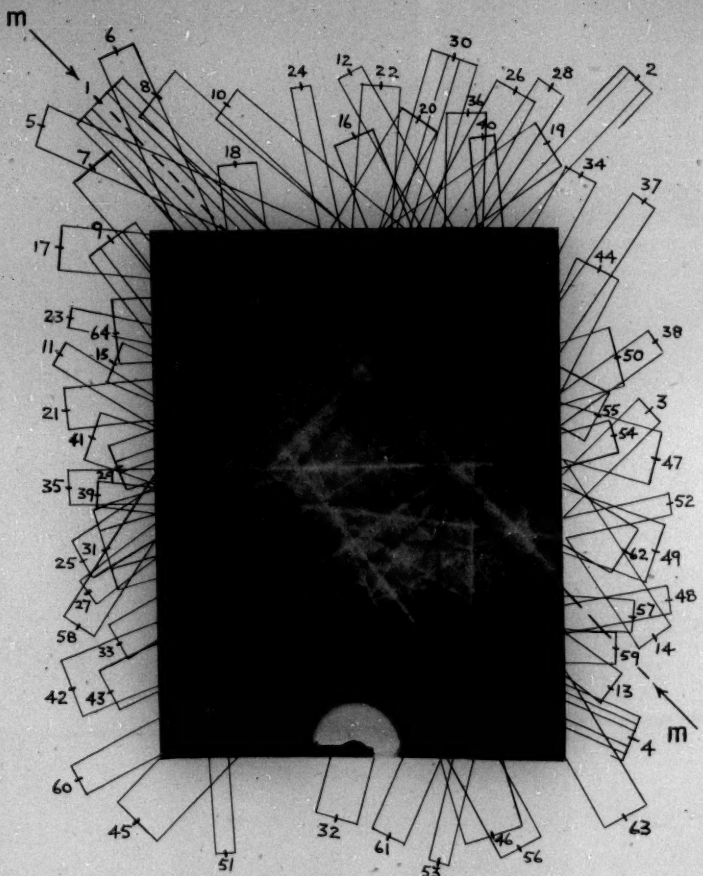


Figure 8 (*a*). Sulphur conchoidal fracture; L~22 cm., 54 kv.; O_1 at intersection of 4 and 37 band medians.

calculated by dividing the above lengths by the scale factor of the drawing or model, and by the estimated λL for the photograph. From the a^*, b^*, c^*, α^*, β^*, γ^* so obtained, the crystal lattice constants a, b, c, α, β, γ can be calculated from the usual relationships. The symmetry of the crystal lattice is of course of the same kind as that of the reciprocal lattice.

Although the determination of the lengths a, b, c involves the value of λL, which is not usually known with great accuracy, the axial ratios and the angles α, β, γ can be obtained fairly accurately as above from the reciprocal lattice measurements, or also from the angles between net-planes by the usual methods used with

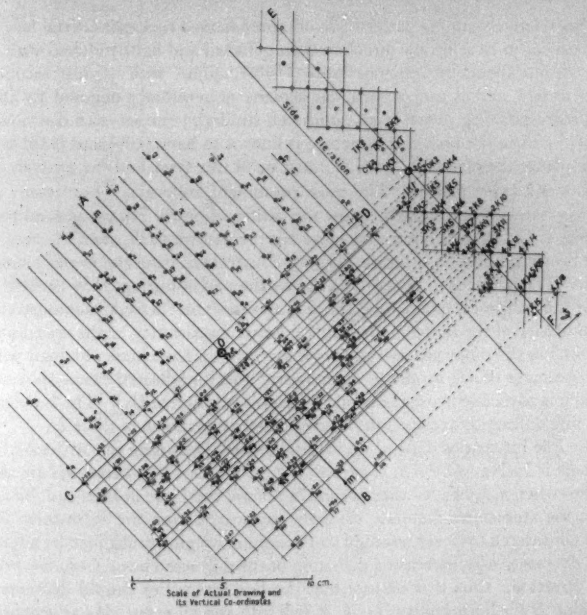


Figure 8 (*b*).

goniometric measurements obtained from crystal faces. In any case, measurement of the angles $(180°—\delta_p)$ enclosing O_1, between suitable line-pair medians in the pattern, can be used to calculate the angles δ between the corresponding net-planes, as check comparisons with the δ values calculated from the lattice constants obtained from the reciprocal-lattice projection. The angles $(180°—\delta_p)$ can usually be estimated to within about 0° 10′ by use of a normal set-square and a transparent rotating-arm set-square having a 4-inch radius scale of degrees. To obtain δ from the measured angle $(180°—\delta_p)$ and the p_1 and p_2 values for the two line-pair medians, we have simply

$$\cos\delta_p = \{\cos\delta - \cos(90°-\phi_1)\cos(90°-\phi_2)\}/\sin(90°-\phi_1)\sin(90°-\phi_2),$$

so that $$\cos\delta = \{\cos\delta_p + (p_1p_2/L^2)\} \,.\, \{1+(p_1/L)^2\}^{-\frac{1}{2}} \,.\, \{1+(p_2/L)^2\}^{-\frac{1}{2}}.$$

§5. EXAMPLES OF RECIPROCAL LATTICE CONSTRUCTION

Two examples are given to illustrate the results and the accuracy obtained in typical cases and to show some features of the line intensities. In these examples, since the camera-length and position of O_1 were not known very precisely, the accuracy obtained in the projections is not the maximum obtainable, but even so, there is no difficulty in recognizing the periodic distribution of the lattice points and the lattice symmetry.

(i) *Rock-salt, NaCl, figure* 7 (*a*)

This relatively simple pattern was obtained from a rock-salt crystal face which had been cut in an arbitrary direction then polished and lightly etched with water. Lines from about fifty planes were visible, some with strong second and higher orders, and as the pattern has a plane of symmetry denoted by the line m–m, corresponding planes were numbered similarly, one set with distinguishing dashes. (Since the rock-salt structure is known to have (100) and (110) types of mirror planes the plane "1", i.e. m, must be of this type, and the analysis shows that it was a $(1\bar{1}0)$ plane.) The most prominent zone axis where many of the narrower bands cross, slightly below the centre of figure 7 (*a*), was used for convenience as O_1, though not a very close approximation to O_1, and the reciprocal-lattice projection shown in figure 7 (*b*) was constructed from the D and p measurements. The projection, figure 7 (*b*), shows the neighbouring body-centred cubic cells of the reciprocal lattice, and the side elevation on the right demonstrates the rectangularity of the lattice. Probably a better consistency of the coordinates of the points in the outer parts of the projection would have been obtained with the closer estimate of O_1 at about 2 cm. nearer the undeflected-beam spot, and in general it is advisable to use the actual estimated position of O_1. The Laue indices of the line diffractions relative to the axes chosen are shown in table 2.

The line intensities depend on various factors, including the structure factor, S^2, which is $16(\mathrm{Na}+\mathrm{Cl})^2$ if h, k, l are all even, $16(\mathrm{Na}-\mathrm{Cl})^2$ when h, k, l are all odd, and zero when h, k, l have other values; h, k, l are the Laue indices and Na and Cl denote the atom-form factors. As a first approximation the atom-form factors can be considered to be represented by the atomic weights multiplied by a function which decreases with increasing deviation of the diffracted beam from the primary beam direction; thus it is natural that the line intensities should decrease with increasing D for both groups, i.e. h, k, l all odd and all even. As an example, in the orders of diffraction from $11\bar{1}$, $11\bar{1}$ has medium, $22\bar{2}$ strong, $33\bar{3}$ nearly zero, and $44\bar{4}$ medium intensity.

The asymmetrical position of the mirror plane m relative to the primary beam enables the relationship between the line intensities and the inclination of the reflecting planes to the beam to be demonstrated clearly in a single reflection pattern by many pairs of band systems. It is possible in this way to see that the expression for the line intensity must contain a factor which reduces the intensity rather rapidly as the angle between the plane and the primary beam (roughly proportional to the perpendicular from the undeflected-beam spot to the band median) approaches zero. Thus, for example, the difference in intensity of the lines of 8 and 8′, 7 and 7′, 10 and 10′ etc., is not large, while in 3 and 3′, and 4 and 4′ it is larger, and in 5 and 5′, 9 and 9′, and 11 and 11′ it is quite high because the planes 5, 9 and 11′ are so much nearer to the primary beam than 5′, 9′ and 11

respectively. Although the more strongly developed narrow bands almost disappear near the undeflected-beam spot, as in the case of 5, the intensified lines (white in figure 7 (*a*)) of wider bands are often strengthened near the foot of the perpendicular from the undeflected-beam spot as in 24′, 25′ and 26, apparently because of the narrowness of the selective-reflection range (corresponding to the line sharpness) in such cases, and the steep gradient of the background intensity near the undeflected-beam spot.

Table 2. NaCl line indices and intensities in figure 7 (*a*)

Band no.	Laue indices *hkl*	Intensity	Band no.	Laue indices *hkl*	Intensity
1	$2\bar{2}0$	VS	—	—	—
	$4\bar{4}0$	M			
	$6\bar{6}0$	F			
2	$11\bar{1}$	M	—	—	—
	$22\bar{2}$	S			
	$33\bar{3}$	~nil			
	$44\bar{4}$	M			
3	$2\bar{2}4$	MF	3′	$\bar{2}24$	MS
4	$0\bar{4}2$	MS	4′	$\bar{4}02$	VS
	$0\bar{8}4$	VF		$\bar{8}04$	F
5	$1\bar{3}1$	MF	5′	$\bar{3}11$	M
	$2\bar{6}2$	M		$\bar{6}22$	M
6	$44\bar{2}$	S	—	—	—
7	$0\bar{6}2$	MS	7′	$\bar{6}02$	S
8	$42\bar{2}$	S	8′	$24\bar{2}$	S
	$84\bar{4}$	VVF		$48\bar{4}$	VVF
9	$2\bar{8}2$	VF	9′	$\bar{8}22$	MF
10	$4\bar{2}0$	S	10′	$\bar{2}40$	MS
	$8\bar{4}0$	F		$\bar{4}80$	VVF
11	$\bar{8}40$	M	11′	$4\bar{6}0$	VF
12	$4\bar{6}2$	M	12′	$\bar{6}42$	M
13	$2\bar{4}2$	S	13′	$\bar{4}22$	M
14	$82\bar{4}$	F	14′	$28\bar{4}$	F
15	$46\bar{4}$	MF	15′	$64\bar{4}$	MF
16	$31\bar{1}$	MS	16′	$13\bar{1}$	M
	$62\bar{2}$	MS	—	$26\bar{2}$	M
17	$\bar{6}20$	M	—	—	—
18	$3\bar{3}1$	MF	—	—	—
	$6\bar{6}2$	F			
19	$51\bar{3}$	F	19′	$15\bar{3}$	F
20	$\bar{5}31$	F	—	—	—
21	$\bar{8}02$	MF	21′	$0\bar{8}2$	
22	$\bar{5}11$	F	—	—	—
23	$\bar{8}42$	VF	23′	$4\bar{8}2$	
24	$0\bar{6}4$	F	24′	$\bar{6}04$	M
25	$24\bar{4}$	F	25′	$42\bar{4}$	M
26	$64\bar{6}$	F	26′	$46\bar{6}$	F
27	$66\bar{4}$	F	—	—	—

S=strong; M=medium; F=faint; V=very.

A further prominent feature of most Kikuchi-line patterns, conspicuous in figure 7 (*a*), is the peculiar weakening or step-down in intensity of certain sections

of some lines; for example both the white and black lines of 3′ almost disappear between the second-order band of 1 and the adjoining first-order band of 4 (or second-order band of 5), while the line-pair of the symmetrical plane 3 behaves similarly across 1 and 4′ (not 4). Such conditions help to make the parabolic line envelopes more prominent, for example that (concave to the left) formed by the white lines of 3′, 4′ (first appearing order), 5′ (first-order), 1 (first-order), and the corresponding black lines of 5, 4 and 3, and the mirror-symmetrical parabola (concave to the right) formed by the black lines of the first group and the white lines of the second.

(ii) *Sulphur, figure* 8 (*a*)

This pattern was from an irregular conchoidal fracture of a sulphur crystal of about 1 cm. diameter. The bands are nearly all narrow and correspondingly diffuse-edged, but by applying the criteria described above for the Bragg reflection positions it was possible to measure the band widths to an accuracy of 0·1 mm. and hence to construct quite satisfactorily the projection of the reciprocal lattice, figure 8 (*b*). Here again it is easily seen that the pattern has a plane of symmetry m represented by the median of band 1, with respect to the band positions and ntensities; but band 2 represents a plane which is at right angles to 1 but is not a mirror plane.

The lattice-point projections in figure 8 (*b*) lie close to a series of equally-spaced lines parallel to AB. They will not all lie exactly on these lines because the crystal is such that the lattice planes of (0*k*0) type are not quite normal to the plane of projection, i.e. the photographic plate; in this case, it is sufficiently close for indexing purposes to use these lines without adding all the closely neighbouring lines which would strictly be required in the projection. The points are also grouped on or near lines parallel to CD, which represent projections of lattice rows inclined downward from left to right at about $8\frac{1}{3}°$, as the z coordinates of the points show. The side elevation represents the points of intersection of all these lattice rows with a plane through O at right angles to the projection plane (represented in section by EF) and to CD, and shows the centred rectangle distribution decisively. Strictly t he projection of the lattice rows should be done upon a plane which is normal to the rows, not to their projections, but the small inclination of the plane of the illustrated side elevation to this plane introduces little error. The data in figure 8 (*b*) thus show that the Kikuchi-line median positions and the line-pair separations correspond to a face-centred rhombic unit cell, and the ratios $a:b:c$ estimated by calculations based on the measurements of the projection combined with the z coordinates were 0·82 : 1 : 1·87. Using the approximately known λL value of 1·13 A. cm. a, b and c were found to be 10·6, 12·9 and 24·2 A. respectively. This is in good agreement with the previous X-ray results, since Mark and Wigner (1924) found rhombic sulphur to have a face-centred rhombic cell with $a = 10{\cdot}61$ A., $b = 12{\cdot}87$ A., $c = 24{\cdot}56$ A. thus $a:b:c = 0{\cdot}820 : 1 : 1{\cdot}900$, with space group V_h^{24} (Fddd) and 128 atoms per cell; Warren and Burwell (1935) who confirmed this space group and calculated the atom positions, found $a = 10{\cdot}48$ A., $b = 12{\cdot}92$ A., $c = 24{\cdot}55$ A., i.e. $a:b:c = 0{\cdot}811 : 1 : 1{\cdot}900$. Trillat and Oketani (1938) concluded from their electron-diffraction transmission patterns that $a = 5{\cdot}2$ A., $b = 6{\cdot}4$ A., $c = 12{\cdot}2$ A., but these values do not take account of spots which occur in some regions of the patterns and show the true Laue-zone spacings, which correspond to cell dimensions close to those of Mark and Wigner (1924) and Warren and Burwell (1935).

Table 3 shows the Laue indices of the observed lines in figure 8 (*a*) and gives, where available, the x-ray structure factor and spot intensity in the rotation pattern, as listed by Warren and Burwell. Unfortunately only eight diffractions

Table 3. Sulphur line indices and intensities in figure 8 (*a*) and x-ray data

Band no.	Laue indices *hkl*	Int.	X-ray data Struct. factor	Int.	Band no.	Laue indices *hkl*	Int.	X-ray data Struct. factor	Int.
1	040	VS	472	S	29	4,$\bar{4}$,12	M		
	080	MF	3	nil		(1$\bar{1}$3)	MF		
	0,12,0	VF			30	4,4,12	MF		
	0,16,0	VVF				(113)	F		
2	206	VS	440	VS	31	4,$\bar{4}$,10	M		
	4,0,12	F				8,$\bar{8}$,20	F		
3	408	M	427	MS	32	4,4,10	VF		
4	1$\bar{3}$1	M			33	4,$\bar{2}$,10	M		
	2$\bar{6}$2	MS			34	4,2,10	MF		
	3$\bar{9}$3	M			35	3$\bar{5}$7	VS		
	4,$\bar{1}$2,4	O			36	357	S		
5	1$\bar{7}$5	S			37	6,2,14	MS		
6	175	S	403	MS	38	6,$\bar{2}$,14	S		
7	0,10,2	MF			39	4$\bar{8}$8	MS		
8	0,$\bar{1}$0,2	F			40	488	MS		
9	062	S	610	S	41	3$\bar{9}$5	M		
	0,12,4	F			42	7,$\bar{5}$,15	MF		
10	0$\bar{6}$2	MS			43	5,$\bar{3}$,11	M		
	—	—			44	5,3,11	F		
11	1$\bar{5}$1	M			45	6,0,14	M		
12	151	MF			46	3,11,9	F		
13	1$\bar{7}$1	F			47	3,$\bar{1}$1,9	F		
14	171	F	59	VF	48	6,$\bar{6}$,10	M		
15	2$\bar{8}$4	MF			49	3,$\bar{1}$1,7	F		
16	284	MF			50	5,$\bar{5}$,15	F		
17	2$\bar{6}$8	MF			51	135	MS		
18	268	F			52	2$\bar{2}$4	M		
19	4,$\bar{2}$,14	F			53	224	F	122	F
20	4,2,14	VF			54	3$\bar{3}$7	MF		
21	3$\bar{5}$9	MS			55	1$\bar{9}$5	F		
22	359	MF			56	195	F		
23	1$\bar{3}$3	F			57	3$\bar{5}$5	M		
24	133	F			58	3,1,11	F		
25	4,$\bar{2}$,12	F			59	4$\bar{6}$6	MF		
26	4,2,12	F			60	4$\bar{2}$8	M		
27	3$\bar{1}$9	MS			61	428	F		
28	319	MS			62	1$\bar{9}$3	F		
					63	193	F		

are common to both lists, but these are all in good agreement except in the intensities of successive orders from the (010) plane, where (080) is evidently much stronger in the electron-diffraction pattern than in the x-ray pattern, presumably due to appreciable secondary or dynamic diffraction. Figure 8 (*a*) also shows further examples of the great reduction in intensity of bands which are parallel to the primary beam or nearly so.

§6. SUMMARY

The interpretation of electron-diffraction Kikuchi-line patterns from single crystals is developed as a powerful, independent and general means of determining the crystal lattice type, dimensions and orientation and the Laue symmetry of the atomic arrangement. This aim is realized by a construction of the reciprocal lattice by methods which are described and illustrated. The satisfactory accuracy of the method is due to the relatively precise definitions for the location of the Bragg reflection positions, derived and stated here for the first time, which enable the method to be used even when the crystal lattice constants are so large that the bands are narrow and diffuse-edged, as with rhombic sulphur. The patterns shown also illustrate features of the line intensities, especially the reduction in intensity when bands pass near to the undeflected-beam spot.

REFERENCES

BOERSCH, H., 1937, *Phys. Z.*, **38**, 1000.
BRANDENBERGER, E., 1937, *Z. Kristallogr.*, **97**, 476.
EMSLIE, A. G., 1934, *Phys. Rev.*, **45**, 43.
FINCH, G. I., QUARRELL, A. J., and WILMAN, H., 1935, *Trans. Faraday Soc.*, **31**, 1051.
FINCH, G. I., and WILMAN, H., 1936, *Proc. Roy. Soc.* A, **155**, 345.
FINCH, G. I., and WILMAN, H., 1937 a, *Trans. Faraday Soc.*, **33**, 1435.
FINCH, G. I., and WILMAN, H., 1937 b, *Ergebn. exakt. Naturw.*, **16**, 353.
HAYASI, T., 1934, *Sci. Rep. Tôhoku Imp. Univ.*, **23**, 491.
KIKUCHI, S., 1928, *Jap. J. Phys.*, **5**, 83 ; see also Nishikawa and Kikuchi.
KIRCHNER, F., 1932, *Ergebn. exakt. Naturw.*, **11**, 64.
LAMLA, E., 1938, *Ann. Phys., Lpz.*, **32**, 178.
VON LAUE, M., 1935, *Ann. Phys., Lpz.*, **23**, 705.
VON LAUE, M., 1936 a, *Ann. Phys., Lpz.*, **25**, 569.
VON LAUE, M., 1936 b, *Phys. Z.*, **37**, 544.
MARK H., and WIGNER, 1924, *Z. phys. Chem.*, **111**, 398.
VON MEIBOM, R., and RUPP, E., 1933, *Z. Phys.*, **82**, 690.
NISHIKAWA, S., and KIKUCHI, S., 1928, *Proc. Imp. Acad. Tokyo*, **4**, 475.
RUTHERFORD, E., and ANDRADE, E. N. DA C., 1914, *Phil. Mag.*, **28**, 263.
SHINOHARA, K., 1932 a, b, *Sci. Pap. Inst. Phys. Chem. Res. Tokyo*, **18**, 223, 315 ; 1932 c, *Ibid.*, **20**, 39.
THIRSK, H. R., and WHITMORE, E. J., 1940, *Trans. Faraday Soc.*, **36**, 565.
THOMSON, G. P., and COCHRANE, W., 1939, *Theory and Practice of Electron Diffraction* (London: Macmillan).
TILLMAN, J. R., 1935, *Phil. Mag.*, **14**, 485.
TRILLAT, J. J., and OKETANI, S., 1938, *Z. Kristallogr.*, **98**, 334.
WARREN, B. E., and BURWELL, J. T., 1935, *J. Chem. Phys.*, **3**, 6.

Thermionic Emission Constants and Band Overlap

BY E. P. WOHLFARTH
Physics Department, University of Leeds

Communicated by E. C. Stoner ; *MS. received* 15 *July* 1947

ABSTRACT. A general treatment is given of the effect of band overlap on thermionic emission. The results are applied to nickel, for which the experimental value of the current constant A of 1380 amp. cm^{-2} deg^{-2} differs markedly from the standard value of 120 amp. cm^{-2} deg^{-2}. The case of nickel has recently been discussed by Sun Nien T'ai and Band (1946), who derive an expression for the variation of A with temperature. It is shown

that their relation strictly applies only to " holes " in an isolated *d*-band and is reliable only over a temperature range such that kT/ϵ_{0d} is considerably less than unity, ϵ_{0d} being the unoccupied width of the *d*-band. With an overlapping *s*-band, as in nickel, there is an increasing transfer of electrons from the *d*-band as the temperature rises. The effect is strongly dependent on the ratio $r_0 = \epsilon_{0s}/\epsilon_{0d}$, where ϵ_{0s} is the occupied width of the *s*-band. Expressions are derived and precise calculations carried out for the transfer effect and the variation of A over a range $0{\cdot}0 \leqslant kT/\epsilon_{0d} \leqslant 1{\cdot}0$ (including temperatures up to about 2000° K.), and for r_0 values of 10, 20 and ∞. The last corresponds in effect to an isolated *d*-band. The results of the calculations are shown in figures 2 and 3. A comparison with the experimental value of A is made using a value of ϵ_{0d} estimated from thermal and magnetic data. The satisfactory agreement previously obtained is due to the adoption of a value of the bandwidth which is incompatible with the experimental evidence, and it is shown that the calculated and observed values of A actually differ by a factor of about 6. Various reasons for this discrepancy are discussed, especially the effect of surface impurities.

§1. INTRODUCTION

In a recent paper Sun Nien T'ai and Band (1946) attempted to interpret anomalous values of thermionic emission current constants on the basis of the electron energy band theory of metals. They consider the effect of band overlap in the vicinity of the surface of the Fermi distribution, and show that under certain circumstances this may lead to abnormally high values of the current constant. In the actual treatment, however, no account seems to be taken of the possibility of a transfer of electrons between the two overlapping bands. The relations obtained are simply those for separate single bands, in each of which there is a fixed number of electrons (or "holes") determined by the degree of overlap. The calculation differs essentially from that for a standard single band only in that the energy density of states is much greater and the width of the unoccupied part of the band much smaller than for a "free-electron" band containing a corresponding number of electrons. To the approximation used, the expressions given appear to be correct, but in substituting a numerical value for the *d*-band energy width for nickel, the value used differs by a factor of about 10 from that which the most direct interpretation of the available experimental evidence seems to indicate. The agreement between the observed value of the current constant A for nickel (1380 amp. cm^{-2} deg^{-2} as compared with the standard theoretical value of 120 amp. cm^{-2} deg^{-2}) and that calculated would thus appear to be largely fortuitous. For this reason, and also because the ideas presented have a bearing on many other physical properties of metals, the problem of thermionic emission constants and band overlap is reconsidered in this paper, with special reference to nickel.

Variations in the value of A, the current constant occurring in the Richardson equation, may be due to a temperature dependence of the thermionic work function χ (Fowler and Guggenheim 1939, p. 484). The Richardson equation, written in the form

$$I = A_0(1-\bar{r})T^2 \exp(-\chi/\boldsymbol{k}T), \qquad \ldots\ldots(1.1)$$

gives the variation of the emission current per unit area with temperature, $\bar{r}$ being the reflection coefficient, which will be very small for most clean metal surfaces, and A_0 being a universal constant having a value of 120 amp. cm^{-2} deg^{-2}. Hence if χ_0 is the value of χ at absolute zero, and A the value of the current constant at a temperature T, then

$$I = A(1-\bar{r})T^2 \exp(-\chi_0/\boldsymbol{k}T), \qquad \ldots\ldots(1.2)$$

where

$$A/A_0 = \exp(\chi_0 - \chi)/\boldsymbol{k}T. \qquad \ldots\ldots(1.3)$$

Usually the experimental results are expressed in the form (1.2), with χ_0 constant, so that a decrease of the work function with temperature will be manifested by an increase of A with T. Such a variation will occur for the band form proposed by Mott (1935) and Slater (1936) for nickel. In this assembly the electrons are distributed in two overlapping energy bands, one of which, the "d-band", corresponding roughly to the $3d$ electrons, is very much higher than the other one, the "s-band", which corresponds to the $4s$ electrons, and which will approximate to a "free-electron" distribution (see figure 1). Fermi-Dirac statistics are readily applicable to this assembly if for the relevant portions of each band $\nu(\epsilon)$, the number of distinguishable states per unit energy range, is proportional to the square root of the energy ϵ. The diagram given by Slater on the basis of numerical calculations indicates that this assumption is justifiable for nickel as giving a reasonable approximation (see § 4). For other metals there may be considerably greater deviation from the standard parabolic band form.

For nickel the equilibrium state at absolute zero is such that the number of electrons in the s-band is equal to the number of absent electrons, or "holes", in the d-band. Evidence drawn mainly from a study of the magnetic properties of nickel and its alloys indicates that at absolute zero approximately 0·6 electrons per Ni atom are in the s-band, and 9·4 electrons (i.e. $10-9{\cdot}4=0{\cdot}6$ holes) in the d-band, as shown in figure 1. It is through the presence of these holes in the d-band that most of the thermal and magnetic properties of nickel have received satisfactory interpretation.

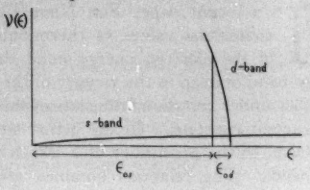


Figure 1. Collective electron assembly for nickel.
ϵ, electronic energy. $\nu(\epsilon)$, energy density of states. ϵ_{0s}, occupied width of the s-band; ϵ_{0d}, unoccupied width of the d-band, both at absolute zero. The diagram is illustrative only (cf. Slater 1936).

ϵ_{0s} and ϵ_{0d} are the respective widths of the relevant portions of s- and d-bands at absolute zero. For nickel

$$\int_0^{\epsilon_{0s}} \nu_s(\epsilon)d\epsilon = \int_{\epsilon_{0s}}^{\epsilon_{0s}+\epsilon_{0d}} \nu_d(\epsilon)d\epsilon \qquad \ldots\ldots(1.4)$$

at all temperatures, and it is seen from the diagram that

$$r_0 = \epsilon_{0s}/\epsilon_{0d} \gg 1. \qquad \ldots\ldots(1.5)$$

As will be shown below the equilibrium appropriate to absolute zero will be displaced as the temperature increases, when there will be a transfer of electrons from the d-band to the s-band. As a result the number of holes in the d-band and the number of electrons in the s-band both increase with a rise in temperature. A method of estimating the extent to which electrons are transferred in this way is given in § 2, and it is shown that the current constant as calculated by considering an isolated narrow d-band has to be modified if the transfer effect is taken into account.

§ 2. THEORY OF THE TRANSFER EFFECT IN NICKEL

In this section is given the quantitative treatment of the electron transfer which takes place in nickel. In the analysis the notation employed is in the main that proposed by Stoner (1939 a), whose paper should also be consulted

for fuller details regarding Fermi-Dirac statistics. The distribution expression applicable to free electrons, or, as here, to the collective electrons in parabolic bands, may be written in the form

$$\frac{dN}{d\epsilon} = \frac{3}{2}\frac{N_0(\epsilon^{\frac{1}{2}}/\epsilon_0^{\frac{3}{2}})}{\exp(\epsilon/\boldsymbol{k}T-\eta)+1}, \qquad \ldots\ldots(2.1)$$

Here N_0 is the number of electrons per atom present in the band and ϵ_0 its occupied width, both at absolute zero. Alternatively N_0 may be the number of holes per atom, in which case ϵ_0 refers to the unoccupied portion of the band. The statistical parameter, η, is related to the chemical potential, or mean Gibbs free energy per particle, μ, by

$$\mu = \boldsymbol{k}T\eta. \qquad \ldots\ldots(2.2)$$

On transforming (2.1), putting $x=\epsilon/\boldsymbol{k}T$, integration over all positive values of x gives

$$N = \tfrac{3}{2}N_0\left(\frac{\boldsymbol{k}T}{\epsilon_0}\right)^{\frac{3}{2}}\int_0^\infty \frac{x^{\frac{1}{2}}dx}{\exp(x-\eta)+1}, \qquad \ldots\ldots(2.3)$$

$$\text{or} \quad N = \tfrac{3}{2}N_0(\boldsymbol{k}T/\epsilon_0)^{\frac{3}{2}}\mathrm{F}(\eta). \qquad \ldots\ldots(2.4)$$

$\mathrm{F}(\eta)$ is the Fermi-Dirac integral given in (2.3). The relation (2.4) states essentially that, as the temperature T increases from absolute zero, the total number of particles per atom also changes, unless

$$\tfrac{3}{2}(\boldsymbol{k}T/\epsilon_0)^{\frac{3}{2}}\mathrm{F}(\eta) = 1. \qquad \ldots\ldots(2.5)$$

This relation holds for a single band, not overlapped by any others. In the present case it is satisfied in the limit $r_0=\epsilon_{0s}/\epsilon_{0d}\to\infty$. The total number of electrons in the band is then necessarily constant. For the assembly to be considered, however, with r_0 finite, the relation (2.5) does not hold for either band.

For the s-band let N_s and N_{0s} be the number of electrons per atom in the band at temperature T and at absolute zero respectively. Also let ϵ_{0s} be the occupied width of the band (cf. figure 1) and $\boldsymbol{k}T\eta_s$ the chemical potential. Then from (2.4)

$$N_s = \tfrac{3}{2}N_{0s}(\boldsymbol{k}T/\epsilon_{0s})^{\frac{3}{2}}\mathrm{F}(\eta_s). \qquad \ldots\ldots(2.6)$$

A similar expression is applicable for the d-band, but here N_d and N_{0d} are the respective numbers of holes per atom. Let ϵ_{0d} be the unoccupied width and $\boldsymbol{k}T\eta_d$ the chemical potential as measured downwards from the top of the d-band. Denoting the chemical potential of the d-electrons, referred to the same level as that of the s-electrons, namely the bottom of the s-band, by $\boldsymbol{k}T\eta'_d$, one obtains

$$\boldsymbol{k}T\eta'_d = (\epsilon_{0s}+\epsilon_{0d})-\boldsymbol{k}T\eta_d. \qquad \ldots\ldots(2.7)$$

The relation corresponding to (2.6) is then

$$N_d = \tfrac{3}{2}N_{0d}(\boldsymbol{k}T/\epsilon_{0d})^{\frac{3}{2}}\mathrm{F}(\eta_d). \qquad \ldots\ldots(2.8)$$

For nickel the condition (1.4), which holds at all temperatures, gives

$$N_d = N_s, \quad N_{0d} = N_{0s}. \qquad \ldots\ldots(2.9)$$

The condition of equality of the chemical potential (i.e. of the Gibbs free energy) for thermodynamic equilibrium gives

$$\eta'_d = \eta_s. \qquad \ldots\ldots(2.10)$$

Insertion of the condition (2.7), (2.9) and (2.10) into (2.6) and (2.8) finally yields the relation

$$\mathrm{F}\{(\epsilon_{0d}+\epsilon_{0s})/\boldsymbol{k}T-\eta_d\}/(\epsilon_{0s})^{\frac{3}{2}}=\mathrm{F}(\eta_d)/(\epsilon_{0d})^{\frac{3}{2}},$$

or, writing $r_0=\epsilon_{0s}/\epsilon_{0d}$ and $\tau_0=\boldsymbol{k}T/\epsilon_{0d}$,

$$\mathrm{F}\{(1+r_0)/\tau_0-\eta_d\}=r_0^{\frac{3}{2}}\mathrm{F}(\eta_d). \qquad \ldots\ldots(2.11)$$

The number of holes in the d-band, N_d, is thus given by (2.8) as

$$N_d/N_{0d}=\tfrac{3}{2}\tau_0^{\frac{3}{2}}\mathrm{F}(\eta_d), \qquad \ldots\ldots(2.12)$$

η_d being determined from (2.11). Owing to the conditions (2.9), the ratio N_s/N_{0s} is equal to N_d/N_{0d} and is thus also given by (2.12). It will be shown below that as T increases the ratio $N_d/N_{0d}=N_s/N_{0s}$ increases from unity at $T=0$. With a rise in temperature the d-band thus loses electrons which are transferred to the s-band. During this process the thermionic work function also changes, and it easy to show from the definition of this quantity (Fowler and Guggenheim 1939, p. 477) that

$$(\chi_0-\chi)=-(\boldsymbol{k}T\eta_d-\epsilon_{0d})=(\boldsymbol{k}T\eta_s-\epsilon_{0s}). \qquad \ldots\ldots(2.13)$$

Hence from (1.3)

$$A/A_0=\exp\,(\tau_0^{-1}-\eta_d). \qquad (2.14)$$

Precise numerical calculations of N_d/N_{0d} and of A/A_0, using (2.12) and (2.14), have been carried out for a series of temperatures (§ 3). These may be usefully supplemented by series expansions appropriate for $\tau_0\ll 1$. For $\tau_0\ll 1$, $\eta_d\gg 1$, $\{(1+r_0)/\tau_0-\eta_d\}\gg 1$, the asymptotic series expansion of the Fermi-Dirac integral may be written in the form

$$\mathrm{F}(\eta_d)=\tfrac{2}{3}\eta_d^{\frac{3}{2}}\{1+(\pi^2/8)\eta_d^{-2}+\ \ldots\}.$$

Similarly, and to the same degree of approximation, writing

$$z=\{(1+r_0)/\tau_0-\eta_d\},$$
$$\mathrm{F}(z)=\tfrac{2}{3}z^{\frac{3}{2}}\{1+(\pi^2/8)z^{-2}\}.$$

Hence from (2.11), after some rearranging,

$$\tau_0^{-1}-\eta_d=\frac{\pi^2}{12}\left[\frac{r_0/\tau_0-\eta_d}{\eta_d\{(1+r_0)/\tau_0-\eta_d\}}\right]. \qquad \ldots\ldots(2.15)$$

To a first approximation the right hand side of (2.15) is zero, so that $\eta_d=\tau_0^{-1}=\epsilon_{0d}/\boldsymbol{k}T$. Inserting this value into the right hand side the second approximation is obtained in the form

$$\tau_0^{-1}-\eta_d=(\pi^2\tau_0/12)(r_0-1)/r_0. \qquad \ldots\ldots(2.16)$$

Hence from (2.8) and (2.9)

$$N_d/N_{0d}=N_s/N_{0s}=1+(\pi^2/8r_0)\tau_0^2. \qquad \ldots\ldots(2.17)$$

The relation (2.17) shows that for $\tau_0>0$, $N_d>N_{0d}$ and $N_s>N_{0s}$, so that there is in fact a transfer of electrons from one band to the other. The temperature variation of the current constant is given by (2.14) and (2.16) as

$$A/A_0=\exp\,\{(\pi^2\tau_0/12)(r_0-1)/r_0\}. \qquad \ldots\ldots(2.18)$$

Since τ_0 will in most cases of physical interest rarely exceed unity (see § 4) it is unnecessary to derive any high temperature approximations, applicable if $\tau_0\gg 1$. A point of interest does, however, arise in connection with (2.11). For $\tau_0\to\infty$ this relation defines a certain lower limit of η_d, which is here denoted by y, where

$$\mathrm{F}(-y)=r_0^{\frac{3}{2}}\mathrm{F}(y). \qquad \ldots\ldots(2.19)$$

For $r_0 > 1$, y will be negative and will be a function of r_0 only. The expression (2.14) now shows that the limiting value of A/A_0 (for $\tau_0 \to \infty$) is very simply related to y by

$$A_{\max}/A_0 = \exp|y|, \qquad \ldots\ldots(2.20)$$

$A_{\max}$ being the maximum value of the current constant for any particular value of r_0. Values of y and $A_{\max}/A_0$ have been calculated for $r_0 = 10$ and 20 and for $r_0 \to \infty$ and are given in § 3 as supplementing the curves shown in figure 3.

As regards the relations derived by Sun Nien T'ai and Band (1946) it may be noted that, since $\tau_0 = \boldsymbol{k}T/\epsilon_{0d}$ the low temperature expression (2.18) is exactly equivalent to their relation (13), which, assuming that $r_0 \gg 1$, was further simplified to give

$$A/A_0 = \exp\{(\pi^2/12)(\boldsymbol{k}T/\epsilon_{0d})\}. \qquad \ldots\ldots(2.21)$$

This expression was taken to be the basic relation and was employed for "complete calculations of current–temperature curves". Such a procedure is, however, open to objection. In the first place the low temperature approximations used are reliable only if $\boldsymbol{k}T/\epsilon_{0d} \ll 1$. For nickel, however, $\epsilon_{0d}/\boldsymbol{k}$ is of the order of 10^3 °K. (see § 4), so that, at the higher temperatures used in experimental determinations of the thermionic constants, $\boldsymbol{k}T/\epsilon_{0d}$ may be of the order unity. For this reason precise calculations of the type outlined in § 3 have to be carried out. Furthermore, by stating their relation (14) in the form given, Sun Nien T'ai and Band seem to neglect the effect of the *s*-band altogether. This relation would in fact be obtained by considering the temperature variation of $(\epsilon_0/\boldsymbol{k}T - \eta)$ for the *d*-band by itself, the transfer effect being thus neglected. A second approximation to the Richardson equation had, in fact, already been derived in this way by Fowler and others (see Reimann 1934), although, since particles rather than holes were considered, a decrease of A with temperature had been predicted.

§ 3. COMPUTATIONAL DETAILS AND RESULTS

The equations to be solved are given in § 2 and are here rewritten for convenience:—

$$\mathrm{F}\{(1+r_0)/\tau_0 - \eta_d\} = r_0^{\frac{1}{2}}\mathrm{F}(\eta_d), \qquad \ldots\ldots(3.1)$$

$$N_d/N_{0d} = \tfrac{2}{3}\tau_0^{\frac{3}{2}}\mathrm{F}(\eta_d), \qquad \ldots\ldots(3.2)$$

$$A/A_0 = \exp\{\tau_0^{-1} - \eta_d\}. \qquad \ldots\ldots(3.3)$$

As has been stated the approximations (2.17) and (2.18) are applicable only if $\tau_0 \ll 1$. Since, however, this condition is not satisfied if the temperature is high and if ϵ_{0d} is relatively small, it may be necessary to obtain exact numerical solutions of the equations for a suitable range of τ_0, say $0{\cdot}0 \leqslant \tau_0 \leqslant 1{\cdot}0$, and for one or more values of r_0. For nickel the magnitude of this parameter is rather uncertain. This limitation of the present treatment is discussed further in § 4, where estimates of ϵ_{0s} and ϵ_{0d} are obtained. These indicate that r_0 lies between 10 and 40. The calculations in this section were carried out for two values of r_0, namely 10 and 20. Additional calculations for $r_0 \to \infty$, corresponding to an infinitely wide *s*-band, can be made very simply, and the diagrams showing the results of the three sets of calculations almost certainly cover the physically relevant range.

It would be most convenient to obtain values of N_d/N_{0d} and of A/A_0 for equally spaced values of τ_0 lying in the desired range. For the present purpose

it was found to be simpler, however, to proceed as follows. For a given value of r_0 a trial value of η_d is first chosen, satisfying the condition that $\eta_d > y$, where y is given by (2.19). The magnitude of the corresponding Fermi-Dirac integral $F(\eta_d)$ is then obtained from the tables of McDougall and Stoner (1938), so that $F\{(1+r_0)/\tau_0 - \eta_d\}$ is calculable from (3.1). To obtain values of the argument inverse interpolation in the same tables would be necessary if r_0 were small. For r_0 equal to 10 or 20, however, the left hand side of (3.1) is large, and the value of $\{(1+r_0)/\tau_0 - \eta_d\}$ can be obtained from the inverse asymptotic series, which is discussed fully by McDougall and Stoner. With

$$\alpha = [\tfrac{3}{2}F\{(1+r_0)/\tau_0 - \eta_d\}]^{-\frac{2}{3}} = r_0^{-1}[\tfrac{3}{2}F(\eta_d)]^{-\frac{2}{3}}$$

it is shown that

$$\{(1+r_0)/\tau_0 - \eta_d\} = \alpha^{-1}\{1 - a_2\alpha^2 - a_4\alpha^4 - a_6\alpha^6 - \ldots.\}, \qquad \ldots\ldots(3.4)$$

where $a_2 = 0{\cdot}822\,467$, $a_4 = 1{\cdot}217\,614$, $a_6 = 9{\cdot}161\,386$.

For the purpose of the present calculations the three terms of the series given in (3.4) were found to give a sufficiently high accuracy. Having obtained $\{(1+r_0)/\tau_0 - \eta_d\}$ from (3.4), η_d is calculated using the same value of r_0, 10 or 20, as before. If this value of τ_0 lies outside the range to be covered ($0{\cdot}0 \leqslant \tau_0 \leqslant 10$), a different value of η_d has to be chosen and the calculation repeated. With a suitable η_d, τ_0 having been obtained by the method outlined, N_d/N_{0d} and A/A_0 are determined from (3.2) and (3.3). About 10 values of τ_0 lying in the desired range were calculated for $r_0 = 10$ and 20 respectively, the computations being carried out with the aid of a Brunsviga calculating machine, using Barlow's Tables (ed. Comrie 1941) for the roots and the Federal Works Agency Tables (1939) for the exponentials. The procedure may be illustrated by giving the details for $r_0 = 10$, $\eta_d = 1{\cdot}0$, the numbers being given to five significant figures. Then

$$\begin{aligned} F(\eta_d) &= 1{\cdot}3964, & \tau_0 &= 0{\cdot}63510, \\ F\{(1+r_0)/\tau_0 - \eta_d\} &= 44{\cdot}157, & N_d/N_{0d} &= 1{\cdot}0601, \\ \{(1+r_0)/\tau_0 - \eta_d\} &= 16{\cdot}320, & A/A_0 &= 1{\cdot}8766. \end{aligned}$$

The final results are most conveniently given in graphical form and are shown in figures 2 and 3. Besides the results for $r_0 = 10$ and 20, those for $r_0 \to \infty$ are also indicated. These were calculated using the relation (2·5) which gives

$$\left.\begin{aligned} \tau_0 &= \{\tfrac{3}{2}F(\eta_d)\}^{-\frac{2}{3}}, \\ N_d/N_{0d} &= 1, \\ A/A_0 &= \exp[\{\tfrac{3}{2}F(\eta_d)\}^{\frac{2}{3}} - \eta_d] \end{aligned}\right\}. \qquad \ldots\ldots(3.5)$$

Figure 2 shows the variation of N_d/N_{0d} with temperature. As is to be expected from the relations (2.5) and (3.2), the transfer effect will decrease in magnitude as r_0 increases, until, when $r_0 \to \infty$, the relation (2.5) is obeyed exactly so that N_d/N_{0d} is unity at all temperatures. The physical explanation is that, since the number of electrons per atom in the s-band at absolute zero remains constant at 0·6, an increase of ϵ_{0s} relative to ϵ_{0d} implies a corresponding decrease in the height $\nu_s(\epsilon_{0s})$, so that it becomes increasingly more difficult for a d-electron to enter the s-band. For the smaller values of r_0 shown in the diagram the electron transfer may become quite considerable. For $r_0 = 10$ and $\tau_0 = 1{\cdot}0$, to consider an extreme example, N_d/N_{0d} is as much as 1·14, so that the number of holes per atom at this temperature will have increased from 0·6 to nearly 0·7.

The possible applications of this effect to other physical properties have been discussed elsewhere.

In figure 3 the three curves give the variation of A/A_0 with temperature for the values of r_0 indicated. It will be noticed that A/A_0 is much less dependent on the value of r_0 than is N_d/N_{0d}, and, furthermore, that for any given value of τ_0, A/A_0 increases with r_0, reaching a maximum when $r_0 \to \infty$, i.e. if the d-band is overlapped by an infinitely wide and low s-band. In that case the increase of the current constant is due entirely to a normal decrease, with temperature, of the work function as given by the relations (2.5) and (2.13). For finite values of r_0, however, the transfer of electrons from the d-band to the s-band causes a relative increase of the work function owing to the tendency of the Fermi limit

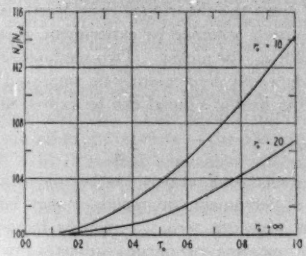


Figure 2. Transfer effect for nickel.

τ_0, reduced temperature kT/ϵ_{0d}. Nd, N_{0d}, number of holes in d-band at reduced temperature τ_0 and at absolute zero respectively. The curve for $r_0 = \infty$ coincides with the τ_0-axis.

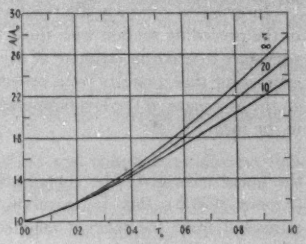


Figure 3. Temperature variation of A/A_0 for nickel.

τ_0, reduced temperature kT/ϵ_{0d}. A, A_0, thermionic emission current constant at reduced temperature τ_0 and at absolute zero respectively. $A_0 = 120$ amp. cm^{-2} deg^{-2}.

to shift to regions of lower energy (see figure 1). Hence the current constant will decrease relative to its particular value for $r_0 \to \infty$. In the extreme case of $r_0 = 1$ the increase in the work function due to the transfer effect just balances the normal temperature decrease, and $A/A_0 = 1$ at all temperatures. This is at once apparent from the relation (3.1), which gives, with $r_0 = 1$, $\eta_d = \tau_0^{-1}$, so that, from (3.3), $A/A_0 = 1$. As has been stated in § 2 an increase of the current constant will occur only if holes are relevant, while for particles A/A_0 will decrease with temperature. For any finite value of r_0, A/A_0 will tend asymptotically to a limiting value $A_{\max}/A_0$ as $\tau_0 \to \infty$, as is indicated by the discussion given in § 2, where it was shown (equation (2.20)) that $A_{\max}/A_0 = \exp|y|$. Calculations of y, using (2.19) can be readily carried out by employing a modified Taylor's series method, the required derivatives of $F(\eta)$ being given by McDougall and Stoner (1938). It is found that for

$$r_0 = 10, \quad y = -2{\cdot}2527, \quad A_{\max}/A_0 = 9{\cdot}5134$$

and for

$$r_0 = 20, \quad y = -2{\cdot}9844, \quad A_{\max}/A_0 = 19{\cdot}775.$$

These figures indicate that $A_{\max}/A_0$ increases with r_0, until, for $r_0 \to \infty$, $A_{\max}/A_0 \to \infty$.

It is of interest in this connection to compare the results obtained by the methods of this section with those obtainable from the low temperature expression (2.18). Selecting a particular value of τ_0, say $\tau_0 = 1{\cdot}0$, the three values of A/A_0

determined roughly from the curves of figure 3 are 2·35 ($\tau_0=10$), 2·56 ($\tau_0=20$) and 2·79 ($\tau_0\to\infty$), while the corresponding values as calculated from (2.18) are 2·10, 2·18, and 2·28 respectively. The magnitude of A/A_0 given by (2.21) (equivalent to equation (14) of Sun Nien T'ai and Band's paper) is equal to that just given for $\tau_0\to\infty$, i.e. $A/A_0=2{\cdot}28$. The discrepancy between the two values of A/A_0 as calculated by the exact methods of this section and from the approximate relation is thus not very large, even for $\tau_0=1{\cdot}0$. The considerations given in the preceding section show, however, that use of the approximation is nevertheless not justified, except when $\tau_0\ll 1$.

§4. COMPARISON WITH EXPERIMENT: CONCLUSION

In order to compare the results deduced from the theoretical considerations of the preceding section with those that have been obtained by experiment, it is necessary to make an estimate of the bandwidth ϵ_{0d} for nickel. Two methods of estimating this quantity from observational data may conveniently be employed. In the first place, if the low temperature specific heat of a metal can be expressed in the form

$$c_v=\gamma T+bT^3,$$

the linear term gives the contribution of the electrons, and it has been shown (Stoner 1939 b), that for parabolic bands in ferromagnetic metals γ may be estimated from the relative magnetization at absolute zero, ζ_0, and the bandwidth ϵ_{0d}. For nickel a consideration of the magnetic properties indicates that ζ_0 is very nearly equal to unity, so that from the experimental value of γ, namely $17{\cdot}44\,.\,10^{-4}$ cals deg^{-2} mol^{-1} (Keesom and Clark 1935), $\epsilon_{0d}/\boldsymbol{k}$ is found to be about 2200° K., provided account is taken of the small contribution of the *s*-electrons. Alternatively $\epsilon_{0d}/\boldsymbol{k}$ may be estimated from the observed Curie temperature, using the relation (Stoner 1938)

$$\epsilon_{0d}/\boldsymbol{k}=\theta_c/f(\zeta_0),$$

applicable to parabolic bands, for which $f(1{\cdot}0)=0{\cdot}41$. Hence with $\theta_c=631°$ K., $\epsilon_{0d}/\boldsymbol{k}$ is found to be about 1600° K. The value of $\epsilon_{0d}/\boldsymbol{k}$ may be estimated on the assumption that the *s*-electrons are "free", as in copper, giving $\epsilon_{0s}/\boldsymbol{k}=6\,.\,10^4$ °K., which has to be regarded as an upper limit. It is difficult to choose between the two estimates of $\epsilon_{0d}/\boldsymbol{k}$ given above and also to see how far the premises underlying the calculation of $\epsilon_{0s}/\boldsymbol{k}$ are correct. The main difficulty appears to be the uncertainty as to deviations of actual band shapes from that postulated, and as to the different effects such deviations may have on the various theoretically determined magnetic and energetic properties of metals. For the electron assembly considered both the *d*-band and the *s*-band may conceivably deviate from a parabolic shape, and the extent of the deviations could only be gauged if the calculations of Slater (1936) were to be carried further with an aim to obtaining a higher accuracy in the final results.

It is fairly certain, however, that the estimate of $\epsilon_{0d}/\boldsymbol{k}$ given by Sun Nien T'ai and Band is far too low. In their paper it is stated that "in fact electronic specific heat data give ϵ_{0d} as small as 0·025 ev.", corresponding to $\epsilon_{0d}/\boldsymbol{k}=300°$ K. The calculation outlined above, on the other hand, shows that the bandwidth deducible from the low temperature specific heat is 2200° K. Using their value of $\epsilon_{0d}/\boldsymbol{k}$, giving, for $T=1400°$ K., $\tau_0=4{\cdot}7$, Sun Nien T'ai and Band were able to explain the anomaly in the observed current constant for nickel quite well. However,

not only is the estimate of the bandwidth erroneous, but the low temperature expression for A/A_0 will certainly not be applicable for $\tau_0 \simeq 5$. It is easy to see in fact that the agreement between theory and experiment is much more unsatisfactory. The most recent observed value of A is 1380 amp. cm^{-2} deg^{-2} (Fox and Bowie 1933), the experimental results used for its calculation being obtained at a mean temperature of about 1400° K. Assuming $\epsilon_{0d}/k = 2000°$ K., i.e. $\tau_0 = 0{\cdot}7$, figure 3 shows that A/A_0 will be of the order of 2·0, i.e. $A \simeq 240$ amp. cm^{-2} deg^{-2}. This discrepancy may be due to a number of causes, a summary of which has been given by Reimann (1934). Among these are:

(i) Surface charging effects. Theoretical discussions of these effects have been given in a number of papers, references to which may be found in Reimann's book. Since, however, the electrons emitted by hot bodies originate in the interior of the emitter, surface effects of this nature may be assumed to be of secondary importance only (see Reimann, p. 271).

(ii) Variation of emitter area. Since any metal surface will in general not be perfectly smooth the actual emitter area will be larger than that calculated geometrically. This would imply an increase of the current constant above the value appropriate to the observed area.

(iii) Thermal expansion effects. These are too complex to discuss fully and reference should be made to the original papers.

(iv) The main reason for the discrepancy appears almost certainly to be the imperfect purity of the emitter surface. As is well known even monatomic impurity layers may influence the thermionic emission very appreciably. This is borne out by the thermionic properties of platinum, which were first investigated by Du Bridge (1928) who found a value of A equal to 17 000. Later work by Van Velzer (1933), who subjected his specimen to some very prolonged heat treatment, showed that for platinum A was very strongly dependent on the presence of gaseous impurities on the surface and in the surrounding atmosphere. The final value of A obtained by Van Velzer was equal to 170, and, it is stated, "slightly lower values of A might have been obtained had the filament withstood further treatment". Fox and Bowie took great care to treat their nickel specimen thoroughly before the investigation. As Van Velzer had shown, however, a very stable state which resisted aging at 1650° K. was attained during the treatment of his platinum specimen. For this state A was abnormally high. Since all the heat treating processes for nickel were carried out at temperatures below 1650° K., the anomalous result for this metal might also be partly indicative of such a state, corresponding to one of imperfect surface purity.

ACKNOWLEDGMENT

The problems discussed in this paper were considered in the course of a more general investigation of the effect of band overlap on the properties of nickel and its alloys with copper, more particularly the thermal and magnetic properties.

The author wishes to express gratitude to Professor E. C. Stoner for his guidance in this and other work.

REFERENCES

Comrie, L. J. (ed.), 1941, *Barlow's Tables* (London : Spon).

Du Bridge, L. A., 1928, *Phys. Rev.*, **32**, 961.

Fowler, R. H., and Guggenheim, E. A., 1939, *Statistical Thermodynamics* (Cambridge : The University Press).

Fox, G. W., and Bowie, R. M., 1933, *Phys. Rev.*, **44**, 345.
Keesom, W. H., and Clark, C. W., 1935, *Physica*, **2**, 513.
Lowan, A. N., *et al.*, 1939, *Tables of the Exponential Function* (New York: Works Project Administration).
McDougall, J., and Stoner, E. C., 1938, *Phil. Trans. Roy. Soc.* A, **237**, 67.
Mott, N. F., 1935, *Proc. Phys. Soc.*, **47**, 571.
Reimann, A. L., 1934, *Thermionic Emission* (London: Chapman and Hall).
Slater, J. C., 1936, *Phys. Rev.*, **49**, 537.
Stoner, E. C., 1938, *Proc. Roy. Soc.* A, **165**, 372; 1939 a, *Phil. Mag.*, **28**, 257; 1939 b, *Proc. Roy. Soc.* A, **169**, 339.
Sun Nien T'ai, and Band, W., 1946, *Proc. Camb. Phil. Soc.*, **42**, 72.
Van Velzer, H. L., 1933, *Phys. Rev.*, **44**, 831.

The Barkhausen Effect.

By H. D. Bush and R. S. Tebble
Department of Physics, The University of Leeds

Communicated by R. Whiddington; *MS. received* 10 *August* 1947

ABSTRACT. A new method of investigating the discontinuous changes in magnetization (the Barkhausen effect) is described. The number of discontinuities occurring is counted, and a knowledge of the change in magnetic moment associated with each one, leads to information concerning the volume of material associated with a discontinuity. The change in magnetization due to discontinuities and the distribution of number and size over the hysteresis cycle is obtained. Conclusions are drawn that the discontinuities associated with volumes of greater than $0{\cdot}9\times10^{-9}$ cm^3 only account for about 30% of the total change in magnetization. Possible explanations of these results are discussed, and suggestions made as to the further development of the method.

§1. Introduction

It is well known that on the steep part of the hysteresis curve for ferromagnetics, the change in magnetization proceeds discontinuously (the Barkhausen effect) and that the discontinuous changes in flux can be amplified as voltage pulses, and displayed on an oscilloscope or made to operate a loud-speaker, when the familiar "rustling" sound can be heard.

The discontinuities have been identified with changes in the directions of magnetization of the domains of ferromagnetic theory. The volume of material constituting a domain and the extent to which the discontinuous processes account for the change in magnetization, are matters of considerable interest, and have been investigated by various observers.

Barkhausen (1919) in his original paper stated that the larger discontinuities corresponded to the reversal of the magnetic moment in volumes which can be estimated to be between 10^{-5} and 10^{-6} cm^3. Tyndall (1924) also found that the larger discontinuities corresponded to a volume of 10^{-6} cm^3, and that the period of the discontinuity was about 10^{-3} seconds.

Bozorth and Dillinger (1929, 1930, 1932) in a series of comprehensive experiments, examined the size of discontinuities in wire specimens, together with their distribution over the hysteresis cycle. They came to the following conclusions: (*a*) The average volume of discontinuities varies over different parts of the field

range, the maximum occurring in the steep part of the hysteresis curve. (*b*) For annealed iron, hard-drawn iron, hard-drawn permalloy and annealed nickel, the discontinuities corresponding to volumes greater than 10^{14} atoms (10^{-9} cm^3), accounted for most of the change in magnetization over the steep part of the hysteresis curve. (*c*) The maximum average volume ranged from $1{\cdot}2 \times 10^{-9}$ cm^3 for annealed iron to $4{\cdot}5 \times 10^{-9}$ cm^3 for malloy. (*d*) It was suggested that these volumes could be identified with the ferromagnetic domains of magnetization.

Bozorth and Dillinger's method was to amplify the voltage pulses produced in a search coil by the discontinuities. From the output current of the amplifier, dB/dt, the rate of change in induction in the specimen caused by discontinuities could be calculated.

Knowing dH/dt, the rate of change of magnetizing field, dB/dH could be found. dB/dH, given by this method of integrating the separate pulses, was compared with the value of dB/dH found from the hysteresis curve.

The average volume of material $\bar{v}$, the magnetization of which changes from saturation in one direction to saturation in the other, is obtained from an expression giving $\bar{v}^2/\bar{v}$ in terms involving, among other factors, the mean square current of the amplifier, the reversible permeability and the conductivity of the specimen.

With the development of modern methods of counting electrical pulses, it was thought that it would be possible to count the number of discontinuities of various sizes which occur when a ferromagnetic is slowly magnetized, so that distribution curves of number against size could be obtained. Further curves could be derived showing the contribution to the magnetization of discontinuities of various sizes, and also the total change in magnetization caused by the total discontinuities detected. This appeared to be a very direct method of approach to the problem.

In this preliminary paper, a description is given of the apparatus developed for this purpose, together with some typical results. The work shows that the method is entirely practicable, and interesting results have already been obtained. The character of the results indicates the directions in which more detailed measurements are required and the developments of the method which are desirable. The results of more comprehensive investigations will, it is hoped, be presented in due course.

§ 2. APPARATUS

A block diagram of the apparatus is shown in figure 1. The two search coils are connected in series opposition and the specimen is placed inside one, this assembly being enclosed in the magnetizing coil. By altering the current in the magnetizing coil by means of a current changing device, the magnetizing field can be slowly changed at an approximately uniform rate.

The changes in flux caused by Barkhausen discontinuities are picked up by the appropriate search coil, and amplified as voltage pulses by the amplifier. The discriminator selects pulses of amplitude greater than any desired value, and these are then counted by a scale-of-sixteen unit and a mechanical recording counter.

The characteristics of the apparatus are as follows. The magnetic field changing unit enables the current in the magnetizing coil to be slowly changed. This is done by charging and discharging a condenser, arranged to alter the grid bias of a pentode valve, thus causing the anode current to change at the required rate. The unit is based on a circuit described by Bozorth (1929). The field can be changed at a rate of 2 to 20 milli-oersteds per second for a maximum field of

20 oersted. For smaller maximum fields the rate of change is proportionately less; in the case of hard-drawn mumetal it was 0·15 to 1·5 milli-oersted per second for a maximum field of 2·6 oersted.

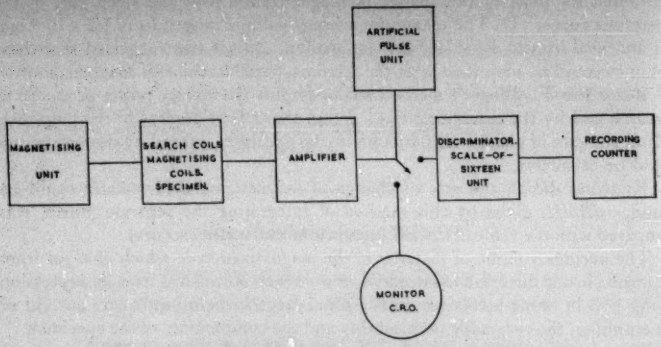


Figure 1. Block diagram of the apparatus.

The amplifier is of conventional design, consisting of five stages of resistance capacity coupling. The noise level with the search coils connected was about 2·5 μv., the first valve being carefully chosen to generate as little noise as possible. The frequency response is shown in figure 2.

The discriminator is based on a design due to Lewis (1942) and selects pulses above any desired size, enabling the various magnitudes of change in magnetic moment, corresponding to Barkhausen discontinuities, to be measured. The discriminator would differentiate between pulses whose amplitudes differed by only 0·5 volts. The scale-of-sixteen unit is built up from the hard valve scale of two circuits originally described by Stevenson and Getting (1937). This scaling unit enabled the actual counting to be done by a recording counter of the watch type (Neher 1939), which had a minimum resolving time of about 1/300 second. A special driving unit was used to operate the recording counter.

An artificial pulse unit was constructed which gave square pulses of about the same time duration as those produced by Barkhausen discontinuities; this was used to calibrate the amplifier and to test the counting circuits. It is of conventional design; the pulse recurrence frequency is determined by a resistance capacity phase shift oscillator, the sine-wave output being fed into a squaring valve. The square wave so produced is passed into a circuit of short time constant for differentiation. The negative and positive pulses produced are passed into a third valve which removes the positive pulse by grid current action, and amplifies and squares the negative one to give a positive pulse output. Variation of the short time constant for differentiation provides a control of pulse duration. The low impedance output is provided by a cathode follower. The

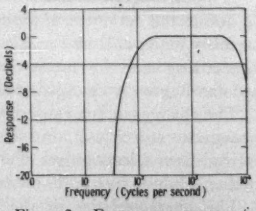


Figure 2. Frequency response of amplifier.

pulse recurrence frequency is variable from 100 to 200 cycles per second and the pulse duration has a range of 300 to 1000 microseconds.

The apparatus can detect changes in magnetic moment of greater than 3×10^{-6} E.M.U. Thus if these changes correspond to complete reversals of the direction of magnetization of domains, the apparatus will detect domain volumes greater than $0{\cdot}9 \times 10^{-9}$ cm^3 in the case of iron where the saturation intensity I_0 is about $1{\cdot}7 \times 10^3$. The domain volume is given by the equation:

$$2I_0 = \Delta M/V,$$

where ΔM is the change in magnetic moment in a volume V. Pulses can be counted at a maximum speed of 5000 per second.

The essential limitation to the sensitivity of the apparatus is set by the noise generated in the amplifier, a large part of which is caused by thermal agitation noise set up by the resistance of the search coils. Hence in the design of the search coils it is important to keep the resistance as low as possible, consistent with the necessity of obtaining the maximum sensitivity. Each search coil consists of 50 000 turns of 44 S.W.G. copper wire, wound in the form of a multi-layer solenoid, 10 cm. long, 0·68 cm. internal diameter and 1·63 cm. external diameter. With these dimensions

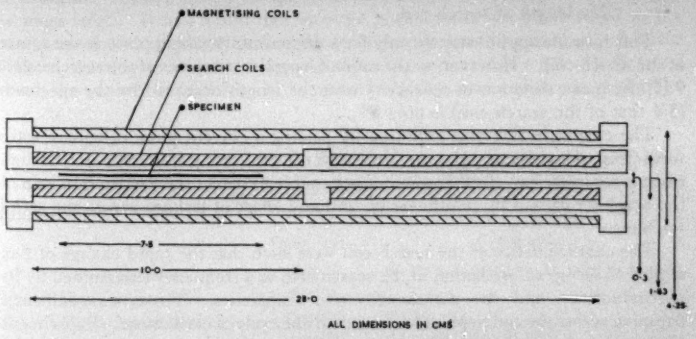


Figure 3. Layout of specimen, search coils, and magnetizing coil.

the response to the changes in flux over most of the length is sensibly constant. Two search coils in series opposition are used to reduce the effect of external interference. Figure 3 shows a section through the assembly of search coils, magnetizing coil, and specimen. The assembly is enclosed in a mumetal box which also reduces the effects of external interference and the effect of stray magnetic fields. With the specimen in the form of a wire 7·5 cm. long placed in the position indicated in figure 3 only one search coil is affected by Barkhausen discontinuities taking place in the specimen.

A fuller description of the apparatus is given in a thesis by one of us (H.D.B. 1947).

§3. CALIBRATION OF THE APPARATUS

The time duration of a Barkhausen reversal, as seen in the output of the amplifier, is of the order of 300 microseconds, and depends on the time constant of the search coil and amplifier, and on the rate of decay of eddy currents in the

specimen. However, the actual change in magnetic moment, and resultant initial change in flux, take place in a time which is short compared with the time constant of the apparatus. Under these last conditions, the amplitude of the resultant output pulse, in the case of a linear amplifier, should be proportional to the change in flux in the search coil.

A calibrating coil, 5 cm. diameter, 10 cm. long and two turns per cm., was placed over the search coil which normally contained the specimen. The flux change produced in the search coil by a known current passing though the calibrating coil could be calculated, making allowance for the variation in field along the length of the coil and over the cross section of the search coil. The relation between flux change in the search coil, and the change in magnetic moment resolved in a direction parallel to the applied field, for a discontinuity taking place in the specimen, is given by:

$$\text{Flux change} = 2\pi\Delta MN/(r^2+l^2)^{\frac{1}{2}},$$

where ΔM = change in magnetic moment;

N = number of turns in search coil;

r = effective radius of search coil;

$2l$ = length of search coil.

This formula applies strictly only for a discontinuity taking place at the centre of the search coil. However, as the ratio of length to diameter of the search coil is 9:1, the mean variation in sensitivity over the length occupied by the specimen (3/4 that of the search coil) is only 1%.

The change in flux was produced by passing a current pulse of square wave form (from the artificial pulse unit), through the calibrating coil, with various high resistances in series. In this way it was possible to detect the variation in gain of the amplifier caused by non-linearity, and any effect of leakage across the series resistance.

The characteristics of the search coil were such that the rapid change of flux tended to set up an oscillation in the search coil, at a frequency determined by its self-inductance and the associated circuit capacities. There was sufficient damping across the coil to prevent more than one cycle of oscillation. Experiment showed that the amplitude of the voltage pulse so produced was proportional to the change in flux, but independent of its duration over a time range of three to one (300 to 1000 μsec.). Pulses from the amplifier were fed into the discriminator, and the setting of the discriminator required to cut off pulses of a given size determined. In this way it was possible to obtain a calibration curve showing variation of change in magnetic moment with the corresponding discriminator setting.

The counting system was tested for accuracy and speed of counting by means of the pulse unit; with sinusoidal pulses, counting speeds up to 5000 cycles per second are possible.

§4. EXPERIMENTAL METHOD

The discriminator was set at a given level; the specimen was then magnetized to saturation in one direction and the field gradually reduced, reversed and then increased to saturation in the opposite direction. Readings of the recording counter were taken, in the case of iron at intervals of 2 oersteds, and in the case of other specimens at other suitable intervals. This was repeated a number of times

at each level of the discriminator and an average taken in each case. A curve was thus obtained for each specimen giving the number of discontinuities N above a given minimum size ΔM plotted against size, as in figure 8 for annealed iron. By differentiating this curve it was possible to obtain a size/number (ΔM vs $dN/d\Delta M$) distribution curve and hence a curve giving size, ΔM, plotted against contribution, $\Delta M . dN/d\Delta M$; see figures 4, 6, 8 and 5, 7, 9 respectively.

The area under the contribution curve

$$\int \Delta M . \frac{dN}{d(\Delta M)} . d(\Delta M),$$

with appropriate limits, is the total change in magnetic moment contributed by discontinuities which can be detected above the noise level of the amplifier. This total divided by the volume of the specimen gives the change in intensity, contributed by the Barkhausen discontinuities measured, over the part of the hysteresis cycle considered.

When the discriminator is set near the noise level there is a background count, constant for each setting, which is appreciable and increases as the noise level is approached. If this becomes comparable with the true count, the error in counting becomes large. It was found that accurate results could not be obtained with a background count of greater than 10 per second at the recording counter. A lower limit to the discriminator setting was therefore put at $\Delta M = 3{\cdot}0 \times 10^{-6}$ E.M.U.; above the setting $\Delta M = 3{\cdot}75 \times 10^{-6}$ E.M.U. the background count was negligible.

§ 5. RESULTS

The following materials, in the form of thin wires 7·5 cm. long, were examined:—

	Diam. cm.	Composition
Annealed iron	0·0376	99·92% Fe, 0·03% C, 0·05% Mn.
Hard-drawn iron	0·0376	
Hard-drawn nickel	0·0458	99·5% Ni.
Hard-drawn mumetal	0·0254	76% Ni, 17% Fe, 5% Cu, 1·5% Cr., 0·6% Mn.

The hysteresis curves for these materials are shown in figures 4 to 7. These curves were obtained using the standard ballistic method, with the specimen in the form of a long wire, and a magnetizing coil 40 cm. long and 1·5 cm. diameter. The change in flux in a search coil of 1000 turns wound round the centre of the specimen, was measured by a Grassot fluxmeter ($1{\cdot}43 \times 10^6$ maxwell turns per radian).

The distribution curves are shown in figures 9, 11, 13. The number of counts per unit range of ΔM plotted as ordinates are reduced to number per cm^3 in order to compare specimens of different sizes. In all cases the curves rise steeply near the noise level.

Size contribution curves are shown in figures 10, 12, and 14, giving in the case of annealed iron a reduction in gradient near the noise level. This reduction may be incorrect because of experimental errors which occur near the noise level, and the errors which are inevitable in differentiation.

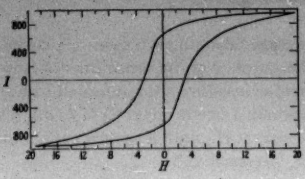


Figure 4. Hysteresis curve for annealed iron.

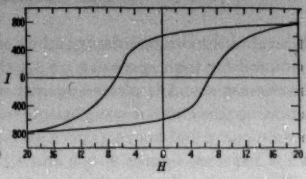


Figure 5. Hysteresis curve for hard-drawn iron.

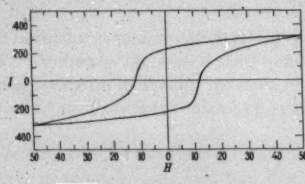


Figure 6. Hysteresis curve for hard-drawn nickel.

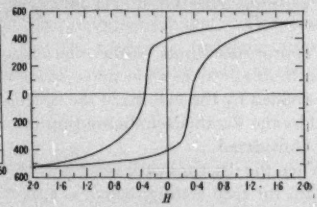


Figure 7. Hysteresis curve for hard-drawn mumetal.

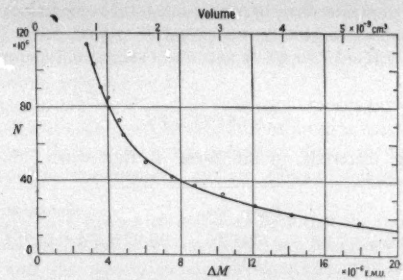


Figure 8. Number-level curve for annealed iron.

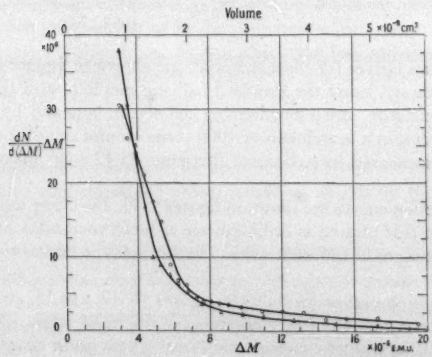


Figure 9. Distribution curves for annealed iron, O, and hard-drawn iron, Δ.

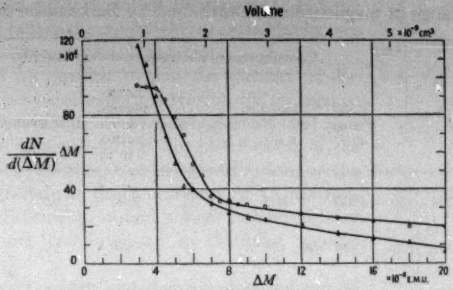


Figure 10. Contribution curves for annealed iron, ○, and hard-drawn iron, Δ.

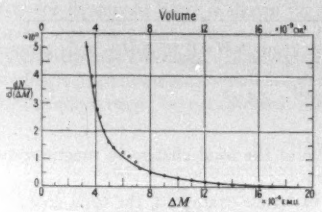


Figure 11. Distribution curve for hard-drawn nickel.

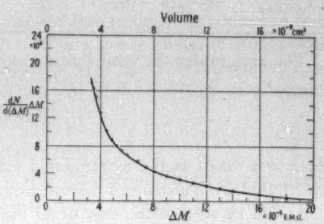


Figure 12. Contribution curve for hard-drawn nickel.

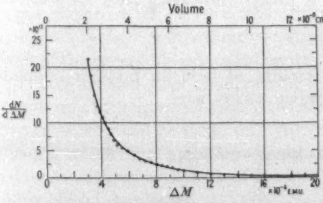


Figure 13. Distribution curve for hard-drawn mumetal.

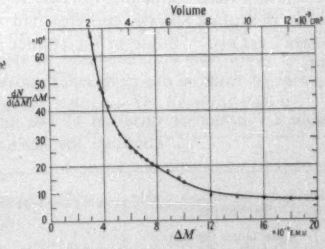


Figure 14. Contribution curve for hard-drawn mumetal.

The contributions to the total change in magnetization which are from the areas under the curves 10, 12, and 14, are given in table 1. The contribution of discontinuities lying above the noise level and below 20×10^{-6} E.M.U. are shown together with the extrapolated contributions. The extrapolated contributions were found in the following manner. The number-level curve over the range which could be examined is clearly of the exponential type. A reasonably satisfactory fit is obtained with a curve of the form $Ae^{-\alpha x^{\frac{1}{2}}}$. Differentiation of this gives the expression $\frac{1}{2}A\alpha x^{-\frac{1}{2}}e^{-\alpha x^{\frac{1}{2}}}$ and the contribution takes the form $\frac{1}{2}A\alpha x^{\frac{1}{2}}e^{-\alpha x^{\frac{1}{2}}}$

Table 1. Change in magnetization contributed by Barkhausen discontinuities

Material	I_0	μ_0	Total change in I	Contribution of Barkhausen discontinuities: From 3 to 20×10^{-6} E.M.U.	% of total	Extrapolated contribution 0 to ∞	% of total	$\frac{dH}{dt}$ 10^{-3} oersted/sec.
Annealed iron	1700	140	1900	605	32	1300	68	3–15
Hard-drawn iron	1700	102	1540	475	31	840	55	3–15
Hard-drawn nickel	500	12·2	520	56	11	120	24	6–30
Hard-drawn mumetal	750	430	1040	320	31	520	50	0·15–1·5

The area under the contribution curve from $\Delta M=0$ to $\Delta M=\infty$ is given by

$$\int_0^\infty \tfrac{1}{2}A\alpha x^{\frac{1}{2}}e^{-\alpha x^{\frac{1}{2}}}=2A/\alpha^2,$$

which was taken as the extrapolated value of the total change in magnetization contributed by Barkhausen discontinuities.

As can be seen in the case of all the materials examined, the total change in magnetization is not accounted for by the Barkhausen discontinuities detected. If this comparison is restricted to the steep portion of the hysteresis cycle the agreement is no better. Further, the rate of change of the magnetizing field was varied within the limits shown, and this had no appreciable effect on the number of counts obtained. Some of the results for nickel, indicating this, are shown in table 2.

Table 2. Effect of variation of rate of change of field on number of counts obtained for hard-drawn nickel wire

Discriminator setting (ΔM) 10^{-6} E.M.U.	3·74	4·14	6·9
Time taken for one half-cycle (minutes)	19, 24, 24, 32, 39	23, 23, 25, 34, 40	27, 27, 30, 60
No. of counts in one half-cycle ($\times10^{-3}$)	106, 119, 109, 103, 110	89, 90, 88, 88, 95	39, 40, 44 40
Mean ($\times10^{-3}$)	110 ± 3	90 ± 2	39 ± 2

With the results obtained it was possible to show the variation in the number of discontinuities per unit range of magnetizing field dN'/dH, where $N'=dN/d\Delta M$ with field H, for different sizes of discontinuities. In order to display these results on a convenient scale, the contribution $\Delta M.dN'/dH$ has been plotted against the field. These curves show maxima near the point $H=H_c$ where H

is the coercive field, as might be expected, and the contribution falls in the region where reversible changes predominate (figures 14, 15 and 16).

It should be possible to plot the number of discontinuities per unit change in intensity, dN'/dI, against intensity I for various sizes of discontinuity and thus obtain some information on the distribution of the discontinuities (number, size and contribution) as the intensity is altered over the hysteresis cycle. Unfortunately it is difficult to obtain a sufficient number of readings where the intensity is changing rapidly (near $H = H_c$).

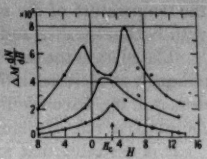


Figure 15. Contribution per unit change in H—annealed iron.

○ $\Delta M = 3{\cdot}2 \times 10^{-6}$ E.M.U.
□ $\Delta M = 6{\cdot}0 \times 10^{-6}$ E.M.U.
△ $\Delta M = 21{\cdot}0 \times 10^{-6}$ E.M.U.

The "average size" of domains detected can be obtained by dividing the total number of discontinuities counted into their contributions for different field ranges. The variation of this "average size" over the hysteresis cycle is shown for annealed iron in figure 18; this is a typical curve, showing a maximum near $H = H_c$. This curve is similar to those obtained by Bozorth and Dillinger. The significance of this curve is limited, as it only applies to those domains which are detected. If all the domains were accounted for the curve might be considerably different.

§6. DISCUSSION

For the materials examined it appears that discontinuous changes in magnetization in domains larger than 0·9 to $3{\cdot}5 \times 10^{-9}$ cm³ according to the material, do not account for the total change in intensity over the steep portion of the hysteresis curve. It would seem that at least some of the discrepancy is caused by domains of a size smaller than those detected by the present apparatus. The extrapolation of the experimental curves gives some estimate of the effect of these smaller discontinuities. Another factor which may account for some of the discrepancy is the extent to which reversible changes in magnetization take place over the hysteresis cycle. A quantitative estimation of this effect can be made by measurement of the reversible permeability at various points on the hysteresis cycle.

There are two ways of increasing the sensitivity of the apparatus so that smaller discontinuities can be detected: (*a*) the noise level of the amplifier could be reduced; (*b*) the sensitivity of the search coil could be increased.

In the case of (*a*) it is doubtful whether the valve noise can be further reduced to any appreciable extent. It does seem possible, however, that the search coils might be re-designed to give a greater sensitivity, without at the same time increasing the thermal agitation noise. This presents difficulties, as increasing the physical dimensions of the search coil will increase the effect of interference from external sources. An additional method might be to reduce

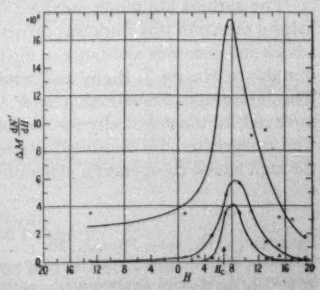


Figure 16. Contribution per unit change in H—hard-drawn iron.

○ $\Delta M = 3{\cdot}2 \times 10^{-6}$ E.M.U.
□ $\Delta M = 6{\cdot}0 \times 10^{-6}$ E.M.U.
△ $\Delta M = 12{\cdot}3 \times 10^{-6}$ E.M.U.

the temperature of the search coils, since the thermal agitation noise decreases with both resistance and temperature.

The effect of the rate of change of field could be investigated over a wider range, although an increase in this rate would appear to increase the possibility of coincidences in the discontinuities. The reduction in size of the specimen would reduce the effective rate of change of field by reducing the number of domains counted.

It is desirable that more results should be obtained near the region $H=H_c$, possibly by increasing the demagnetizing factor of the specimen, so as to reduce the effective gradient of the hysteresis curve in that region. Annealed mumetal would appear to be suitable for this purpose. It would be further interesting to investigate materials of the mumetal type, as it is possible that domain size bears some relation to initial permeability.

It is hoped to carry out experiments on materials covering a wider range of magnetic properties than has been possible so far.

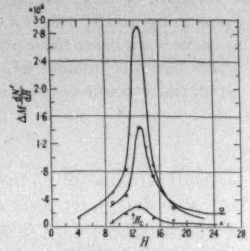


Figure 17. Contribution per unit change in H—hard-drawn nickel.

○ $\Delta M = 3{\cdot}9\times10^{-6}$ E.M.U.
□ $\Delta M = 6{\cdot}2\times10^{-6}$ E.M.U.
△ $\Delta M = 13{\cdot}2\times10^{-6}$ E.M.U.

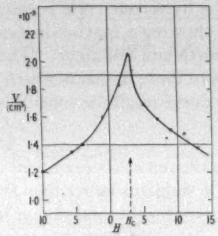


Figure 18. Average domain size for annealed iron.

ACKNOWLEDGMENT

The authors are much indebted to Professor R. Whiddington, under whose helpful direction this work was carried out, and to Professor E. C. Stoner, for his advice and generous assistance in the preparation of the manuscript. Thanks are due also to Messrs. J. Rigby and Sons Ltd., for supplying the iron specimens, to The Telegraph Construction and Maintenance Co. Ltd., who provided the mumetal, to the Admiralty for the loan of apparatus, to Mr. W. A. Ludbrooke, who constructed the counting recorder and search coils, and to Mr. H. Partridge through whose co-operation many of the components were obtained.

REFERENCES

Barkhausen, H., 1919, *Phys. Z.*, **20**, 401.
Bozorth, R. M., 1929, *Phys. Rev.*, **34**, 772.
Bozorth, R. M., and Dillinger, J., 1930, *Phys. Rev.*, **35**, 733 ; 1932, *Ibid.*, **41**, 345.
Bush, H. D., 1947, Ph.D. Thesis, *The Barkhausen Effect*, Leeds University.
Lewis, W. B., 1942, *Electrical Counting* (Cambridge University Press).
Neher, H. V., 1939, *Rev. Sci. Instrum.*, **10**, 162.
Stevenson, E. C., and Getting, I. A., 1937, *Rev. Sci. Instrum.*, **8**, 414.
Tyndall, E. P. T., 1924, *Phys. Rev.*, **24**, 439.

DISCUSSION

Prof. BATES. I should like to emphasize the importance of this work, which I very much hope will be extended. The trouble is that the Barkhausen effect is comparatively easy to show on the lecture table, and people have therefore been inclined to think that everything is known about it. Therefore it is only when a serious and careful study is made that we find, as the authors have clearly shown, things are not at all satisfactory. Moreover, there has been a tendency, particularly in the American and German literature, to give magnetic domains a kind of rigid cubic-block structure that they cannot possibly possess. I very much hope that the extension of this work will enable us to get a clearer view of domain processes, and, in particular, help us to decide on the correctness of the views concerning fillet-shaped domains which have recently been put forward.

Prof. W. SUCKSMITH. Do the authors believe that the discrepancy between the integrated sum of the intensity changes as determined by the Barkhausen discontinuities and the total change of intensity determined in the usual way is due to irreversible groups below the noise level, or to continuous and reversible changes ?

Mr. MCCAIG. I should like to point out that fast counting rates can now be measured without any special form of mechanical counter. Counting speeds of more than 100,000 per second have now been achieved in commerical scaling circuits.

AUTHORS' reply. Measurements on reversible permeability are being carried out at Leeds, the results of which suggest that reversible changes account for about 10% of the total change in magnetization over the range considered. It is possible that some of the discrepancy may be due to the assumption that the time constant of the discontinuity in the specimen is short compared with that of the search coil. This is based on a formula of Wwedensky (1921). It is intended to carry out measurements on the time constant of the discontinuities, making use of a wide-band amplifier (1 kc/s. to 1 Mc/s.) with which it should also be possible to detect the smaller discontinuities.

WWEDENSKY, 1921, *Ann. d. Phys., Lpz.*, **64**, 609.

A Radiosonde Method for Atmospheric Potential Gradient Measurements

BY R. E. BELIN
Dominion Physical Laboratory, Wellington, New Zealand

MS. received 9 *July* 1947

ABSTRACT. A method is outlined whereby the National Bureau of Standards radiosonde, in a modified form, is applied to the measurement of the earth's potential gradient. This paper describes the modification of the sonde circuit to measure the point-discharge current produced by a gradient and the result of a typical sounding of a cumulo-nimbus cloud.

§1. INTRODUCTION

EXPERIMENTS on the variation in natural potential gradient have so far been almost wholly carried out on or near the surface of the earth. The main exceptions were in the work of Simpson and Scrase (1927, 1935), who used a balloon to carry aloft, to heights of the order of 10 km., recording equipment capable of measuring pressure, humidity and potential gradient. The pressure and humidity were recorded, in the usual manner, by employing

an aneroid barometer and a human hair hygrometer respectively; the measurement and recording of the gradient was, however, more complex. Simpson and Scrase relied upon the fact that, if two sharp points with opposed senses are connected by a straight conductor, the axis of which is parallel to a field of suitable magnitude, point discharge will take place, thereby tending to equalize the potentials between the respective parallel planes through the points. In their apparatus this discharge current was caused to flow between two electrodes on polarity paper, and the magnitude of the gradient could be roughly estimated by measuring the width of the line traced by the positive electrode. Using a 20-metre length of fine copper wire to separate the points, it was possible to measure fields not smaller than 10v/cm. and the main disadvantage was that results depended upon the recovery of the equipment intact.

The method employed by the author (1946) differs essentially from that of Simpson and Scrase in that a permanent record is obtained at the receiving station so that results do not rely on the recovery of the sounding equipment. The point-discharge method was used in conjunction with a radiosonde instead of polarity paper. This was so modified that it was capable of transmitting a carrier, the modulation of which was dependent upon the magnitude of the electric field. These signals were detected and recorded by the standard radiosonde receiver developed by the Bureau of Standards (Diamond *et al.* 1940).

§2. METHOD

If two points be placed in an electric field as explained above, and a large-valued resistor be connected in series with the collector wire, at its centre, then the point-discharge current i flowing out of the points will be given (Simpson and Scrase 1927, 1935) by the relation

$$i = a(F^2 - M^2) \qquad \ldots\ldots(1)$$

F is the existing potential gradient (greater than M). M is the minimum gradient required for point discharge to commence. a is a constant depending mainly on the exposure of the points (i.e. pointedness, number of points, distance between them etc.) and to a lesser degree upon the value of the series resistor. This current i produces across the resistor a potential drop which is measured and transmitted by the modified radiosonde.

Thus by carrying aloft (with a neoprene balloon) the collector points separated by a length L of wire with the modified transmitter at its centre the measurement of gradient can be readily achieved.

§3. MODIFIED RADIOSONDE

The radiosonde developed by Diamond, Hinman, Dunmore and Lapham (1940) consisted of two oscillators operating at frequencies of 1 Mc/s. and 70 Mc/s. respectively. The 1-Mc/s. oscillator M (figure 1) was of the squegger type. Its squegging frequency was determined by the R_2C_2 combination introduced into the grid circuit. The oscillator C was also of the squegger type being on when M is off and vice versa, thus effecting a modulation.

When M is oscillating it draws grid current which charges up the condenser C_2 and thus developing across R_2 a potential which builds up and eventually blocks the tube. The period of no oscillation is determined by the time constant R_2C_2. If a bias battery P be introduced, as is done in figure 1, the squegging frequency will be altered. It is found that the squegging frequency has an

approximately linear relationship with the potential developed across R_2 (figure 2). This linear relation holds for positive bias, but only approximately for negative bias up to cut-off. Figure 2 gives steady grid voltage as a function of squegging frequency, when $R_2 = 2\ M\Omega$ and $C_2 = 0{\cdot}003\ \mu F$.

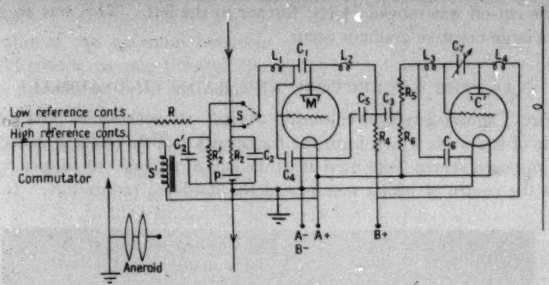


Figure 1. The modified radiosonde circuit.
$R = 1\ M\Omega$, $R_2 = 2\ M\Omega$, $R_2' = 1{\cdot}25\ M\Omega$, $C_2 = {\cdot}005\ \mu F.$, $C_2' = {\cdot}003\ \mu F.$

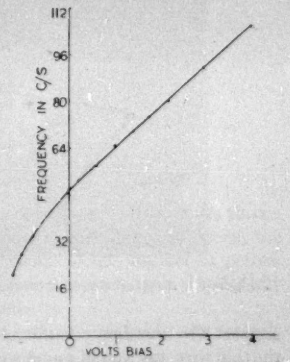


Figure 2.

In figure 1 the small switch S, actuated by the standard relay S′ on the 5, 10, 20, 25 etc. commutator contacts, was used to introduce periodically into the grid circuit of M a fixed time constant R_2' in parallel with C_2'. This periodicity was determined by the rate of change of pressure, or height, and hence was only controlled by the baro-switch. (The baro-switch consisted of a commutator composed of 80 separately insulated silver contacts of which only 5, 10, 15, 20 etc. were used. Over this commutator the contact arm of the aneroid barometer could move freely so that its position was determined by the pressure, that is by the height.) By means of this reference signal, as it was termed, it was possible to evaluate the altitude of the sounding equipment, as well as to obtain a check on any possible oscillator drift. The values of R_2' and C_2' were chosen to produce a squegging frequency of 190 c/s. and were used as the high reference value.

The baro-switch also introduces on the 15, 30, 45 etc. contacts, a resistor R in parallel with R_2. This latter combination produced, when no point discharge was taking place, a low reference of 140 c/s. Thus this low reference was subject to a change, the magnitude and direction of which would be dependent upon the magnitude and sign respectively of the potential gradient. However, the chief function of this reference was not so much a check on the oscillator as a check on the vertical course of the balloon. It served also as a height signal. It is possible should the ascent take place through a cumulo-nimbus cloud, as was intended, to encounter such a strong downward current as to produce a

descent of the balloon. This would produce a change in the predetermined sequence of high and low references which would be observed on the record.

The value of R_2 and C_2 were chosen to give an audio frequency of 50 c/s. with no bias battery in circuit. When the small $1\frac{1}{2}$-volt penlight cell was introduced in series with R_2, a squegging frequency of 64 c/s. was obtained. Thus the cut-off was moved 14 c/s. further to the left. This was an advantage should a large negative gradient occur.

§4. THE COLLECTORS AND RADIATING AERIAL

Figure 3 shows a typical collecting device. It consists of pins soldered to the ends of the wires which form the framework. Two of these collectors with pins in opposed senses were used at either end of a gauge 30 copper connecting wire, at the centre of which was located the modified radiosonde. In order to

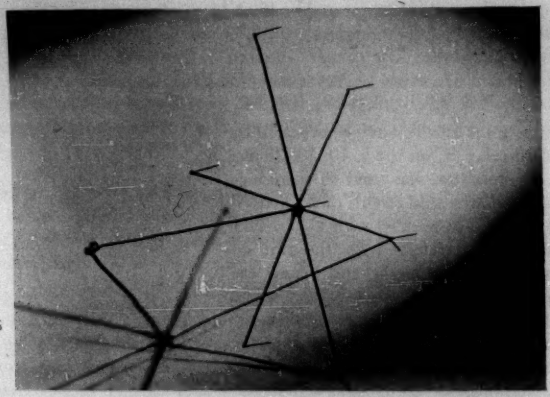


Figure 3. A typical collector, showing the manner in which the sharp points were connected.

support the sounding equipment which includes the collectors and radiosonde, the thin wire was wound round some strong hemp string, while the upper collector was fastened with a similar string, 10 feet long, to the balloon.

The transmitter aerial, a half-wave dipole, was fastened parallel to the collector wires, but separated from them by stand-off insulators of cork (for lightness) impregnated with polystyrene solution in toluene.

The length of the collector wire was chosen according to the formula $(2n-1)$ half-wavelengths (where n is an integer) to give, as it was experimentally shown, a maximum signal strength. The value of n was chosen as 5 which thus gave a length of 64 feet between the collectors for a frequency of 69 Mc/s. The 64-foot connecting wire consisted of gauge 30 enamelled copper wire. For this length the minimum potential gradient capable of beginning point discharge was determined experimentally as 3 v/cm. at N.T.P. Figure 1 shows schematically the collector wire connected to either end of the R_2C_2 combination, while the actual radiating antenna is seen on the right of the circuit.

§5. THE CALIBRATION

Calibration was carried out by means of an adjustable electrostatic field between two large parallel metal sheets. In this field were placed the collectors, actually used, with the pins parallel to the field direction. The collectors were connected by a short wire with the resistor R_2 in series, as the 64-foot length was impracticable.

The value of the potential between two parallel planes passing through the points was determined from the relation

$$\text{potential} = Ldv/dx \text{ volts},$$

where L is the distance in cm. between the pin points and dv/dx the field strength in volts/cm. For points of given shape the value of Ldv/dx was shown to be constant for any value of L. Hence by calibrating this for a small given value of L, the calibration of frequency in cycles/sec. as a function of Ldv/dx was the same for any value of L*. A calibration of a typical sonde using 7 points on both collectors is shown in figure 4. It will be observed that the minimum value of Ldv/dx necessary to initiate point discharge is approximately 6800 volts, corresponding, for the length of collector wire used, to 3·5 v/cm. ; it is thus equal to the value of M required in equation (1), at the existing pressure, temperature and humidity.

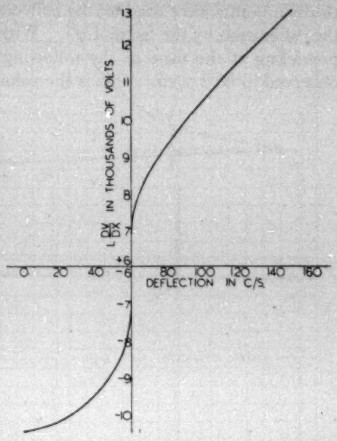


Figure 4.

The minimum value of Ldv/dx necessary to produce a discharge was shown to be independent of the number of points, but once discharge had set in, the point-discharge current became directly proportional to the number of points. Thus when the current for a given number of points is measured, that for any other number of points can be readily determined.

Point-discharge current is also a function of pressure related by the formula (Tamm 1901)

$$i = i_0\{P_0/P - (V/10^5)^{\frac{1}{2}} \log_e(P_0/P)\}^2, \qquad \ldots\ldots(2)$$

where i and i_0 are the currents at pressures P and P_0 respectively and $V = Ldv/dx$. P_0 is the pressure at ground level at time of releasing sounding equipment.

This cumbersome formula may be reduced, without affecting the accuracy required, to the approximate expression,

$$i = i_0 P_0^{1\cdot6}/P, \qquad \ldots\ldots(3)$$

which was used in these experiments.

The height of the sonde was determined from the formula

$$h = 8 \log_e P_0/P; \qquad \ldots\ldots(4)$$

combining equations (3) and (4) we obtain

$$i = i_0 e^{0\cdot2h}.$$

* By increasing the value of L it is possible to make the apparatus as sensitive as is desired.

Hence expression (1) can be modified to include a correction for height

$$i = a(F^2 - M^2)e^{0\cdot 2h}. \qquad \ldots\ldots(5)$$

By the use of this formula the actual value of F, i.e. dv/dx, can be determined at any height.

§ 6. RESULTS

The equipment outlined above was used with only one modification, viz. thirteen points were used on the collector frames instead of seven. This altered the calibration by the factor 13/7. With the temperature, pressure and humidity prevailing at the time of the following ascent, the striking gradient value was observed to be 3 v/cm.; this is the value used in figures 5 and 6.

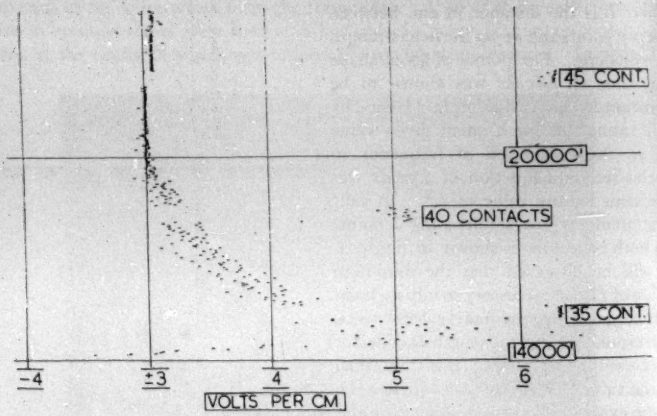


Figure 5.

Figure 6 shows a typical record obtained by releasing this apparatus beneath a cumulo-nimbus cloud with fracto-stratus formation below. The field at ground level was 4 v/cm. positive, that is, downwards. It increased in a fluctuating manner to a maximum of about 5 v/cm. at 600 feet, and then slowly decreased to zero at 2000 feet. The graph has been smoothed particularly in the lowest negative bulge between 2000 and 3500 feet which region probably corresponds to the fracto-stratus region. On entering the cloud at 3500 feet the potential gradient increased rapidly to a high positive value (off record) in the vicinity of 7000 feet. It then began to decrease again and at 10 000 feet suddenly reversed to a high negative value (off record). At this stage the gradient was so large that the discharge current produced a voltage across the grid resistor R_2 which was well beyond the cut-off value. However, a signal was still detected in the form of the 69-Mc/s. carrier wave. At 14 000 feet, the cloud top, the gradient slowly became less negative, swinging over to positive when leaving the cloud, and finally reaching another subsidiary maximum of about 6 v/cm. at 14 600 feet. Above this the gradient decreased again according to the inverse square law until at about 20 000 feet the point discharge current became zero.

A reproduction of a part of the actual record obtained, in this case, when emerging from the cloud top is shown in figure 5, which corresponds to the upper part of figure 6. This record shows, by the scattered points, the fluctuating manne

in which the gradient increases or decreases during an ascent. This fluctuating increase could be explained by slight variations in the state of ionization with height, or possibly a humidity variation which might produce the same result.

Interpreting the curve of figure 6, we note that the cumulo-nimbus appeared to have a large positive distribution of electricity in the base and a large negative distribution in the top. This corresponds to a strong positive gradient below and above the cloud, the latter of which was observed. The former was distorted by the presence of the fracto-stratus formation on which it was assumed the main field of the cloud terminated, thus producing in effect a negative and a positive change in the upper and lower parts of the formation.

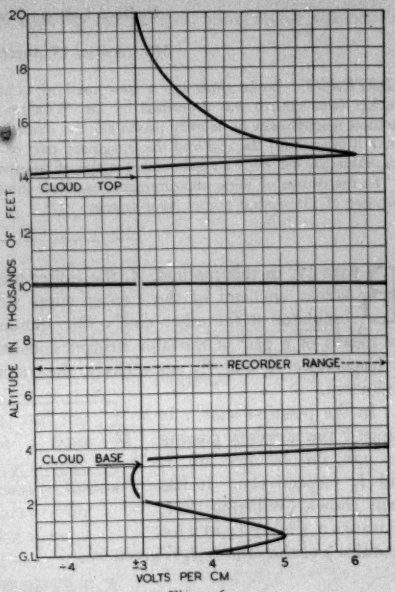


Figure 6.

§7. CONCLUSION

The method outlined above is applicable to many problems relating to potential gradient: it is excellent, as proved, for the measurement of electrical distribution of clouds; it is possible, by lengthening L, to measure the fluctuation in dv/dx with height on a fine day, when it is approximately equal to 1v/cm.; by slight modification to the sonde it is possible to include in the record both humidity and temperature and thus obtain reliable quantitative data on the true variation of the potential gradient; again, by a further slight modification to the circuit which is at present being attempted, it will be possible to obtain measurements of very large gradients, especially negative, which have hitherto been beyond the range of the apparatus.

ACKNOWLEDGMENTS

The author is indebted to Dr. K. Kreielsheimer for valuable assistance and for the loan of his radiosonde receiver, without which these investigations could not have been undertaken, to Professor P. W. Burbidge for helpful suggestions, and to the R.N.Z.A.F. Meteorological Branch for materials, meteorological data and assistance with the balloon flights.

REFERENCES

DIAMOND, HINMAN, DUNMORE, and LAPHAM, 1940, *Bur. Stand. J. Res.*, **25**, 327.
KREIELSHEIMER and BELIN, 1946, *Nature, Lond.*, **157**, 227.
SIMPSON and SCRASE, 1927, *Proc. Roy. Soc.* A, **114**, 376.
SIMPSON and SCRASE, 1935, *Proc. Roy. Soc.* A, **161**, 309.
TAMM, 1901, *Ann. Phys., Lpz.*, **6**, 259.

A Note on the Lecher Wire Method of Measuring Impedance

By MARJORIE WILLIAMSON
Bedford College, London

MS. received 18 *August* 1947

ABSTRACT. Some time ago a new method of measuring impedance by means of a Lecher wire system was described by Dr. Williams in these Proceedings (1944). This method would seem to be very useful for the investigation of the properties of dielectrics since it gives both the real and the imaginary part of the dielectric constant, and it is also adaptable for use over a considerable range of frequencies.

In the course of some measurements on these lines at a wavelength of about a metre a number of interesting points have arisen.

BRIEFLY, in this method, Z, the impedance to be measured, is connected across one end of the Lecher wires which are loosely coupled to an oscillator at the other end. The standing-wave pattern is investigated by means of a pair of movable bridges consisting of vacuum thermojunctions of impedance Z_1 and Z_2 respectively, the values of which need not, however, be known.

In the position shown in figure 1 the currents passing through the bridges are I_1 and I_2 respectively. It has been shown (Williams 1944, equation (10)) that the modulus ρ of the ratio I_1/I_2 is given by

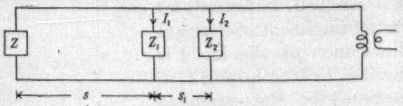


Figure 1.

$$\rho^2 = K_1 + \frac{K_2 + K_3 \sin 2(b+\beta s)}{\sinh^2 a + \sin^2 (b+\beta s)} \qquad \ldots\ldots(1)$$

where $\beta = 2\pi/\lambda$, the current reflection coefficient

$$(Z_0 - Z)/(Z_0 + Z) = e^{-2(a+jb)}, \qquad \ldots\ldots(2)$$

$$K_1 = (1/|Z_1|^2)[(R_2^2 + X_2^2)\cos 2\beta s_1 + X_2 Z_0 \sin 2\beta s_1 + Z_0^2 \sin^2 \beta s_1], \qquad \ldots\ldots(3)$$

$$K_2 = (1/|Z_1|^2)[(R_2^2 + X_2^2)\sin^2 \beta s_1 \cosh 2a + R_2 Z_0 \sin^2 \beta s_1 \sinh 2a], \qquad \ldots\ldots(4)$$

$$K_3 = (1/|Z_1|^2)[(R_2^2 + X_2^2)\sin \beta s_1 \cos \beta s_1 + X_2 Z_0 \sin^2 \beta s_1] \qquad \ldots\ldots(5)$$

and Z_0 is the characteristic impedance of the Lecher wires.

In order to simplify equation (1) the constant K_3 is made zero by adjusting the value of s_1. The necessary condition is, from (5),

$$(Z_2^2 + Z_0 X_2 \tan \beta s_1) \sin \beta s_1 \cos \beta s_1 = 0,$$

i.e.

$$\tan \beta s_1 = -Z_2^2/Z_0 X_2, \qquad \ldots\ldots(6)$$

the other solutions being of no practical importance.

This critical value of s_1, which we shall call s_c, was determined experimentally in the original method. The impedance Z was removed and the lines short

circuited by a copper plate. In this case a and b should be zero and the equation (1) reduces to

$$\rho^2 = K_1 + K_2 \operatorname{cosec}^2 \beta s + 2K_3 \cot \beta s, \qquad \ldots\ldots(7)$$

K_1 and K_3 being unchanged and K_2 simplified by the vanishing of the term with $\sinh 2a$.

If now $K_3 = 0$, a graph of ρ^2 against s is symmetrical about its turning points. A finite value of K_3 introduces an asymmetry and this, in the original paper, is the test for the critical setting.

For $K_3 = 0$ a graph of ρ^2 against $\operatorname{cosec}^2 \beta s$ is a straight line from which K_1, which is necessary for the determination of a, is obtained.

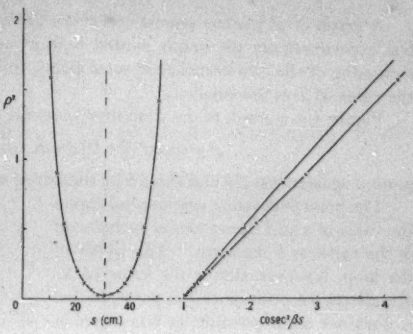


Figure 2.

It was noticed, however, in our experiments, that to a (ρ^2, s) curve which appeared to be symmetrical there corresponded a $(\rho^2, \operatorname{cosec}^2 \beta s)$ graph which was a curve with two branches as shown in figure 2, i.e. K_3 was still finite, and this suggested that there might be a more sensitive test of the accuracy of the setting. It was thought of interest to examine the equation (7) more closely for values of s_1 not equal to s_c in the hope of finding a quick and accurate way of determining s_c.

The form of the curve given by equation (7) is shown in figure 3 which is a graph of the simplified function

$$\rho^2 = \operatorname{cosec}^2 \theta + K \cot \theta$$

from $\theta = 0$ to $\theta = \pi$ for the arbitrary values $K = 1$, $K = \frac{1}{2}$. For values of θ greater than π (i.e. for increasing s) the curve is repeated.

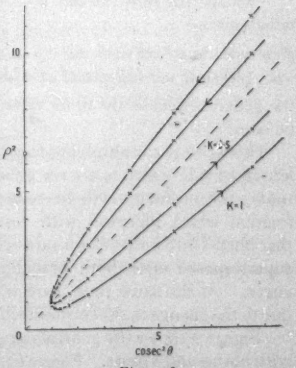


Figure 3.

As θ increases the curve is traversed in the direction shown by the arrows if K is positive, and in the reverse direction if K is negative. Hence from the experimental curves it is quite obvious when the critical setting has been overshot. As s_1 approaches the value s_c, i.e. as K tends to zero, the curves close in and approximate more and more nearly to the required straight line.

A number of experimental curves were obtained for various values of s_1, and while the general behaviour corresponded to the above analysis in that the branches of the curves closed in and opened out again on passing through the critical setting, most of the curves showed unexpected peculiarities of curvature,

and in some cases there was a crossing of the two branches. In no case could a reasonably good straight line be obtained.

There is nothing in equation (7) to account for this, but a consideration of the general equation (1) shows that if a or b or both are finite, though small, the (ρ^2, $\mathrm{cosec}^2\,\beta s$) graph is considerably altered.

Take first the case $a=0$, b finite; then

$$\rho^2 = K_1 + K_2\,\mathrm{cosec}^2\,(b+\beta s) + 2K_3\,\cot\,(b+\beta s).$$

A graph of ρ^2 plotted against $\mathrm{cosec}^2(b+\beta s)$ would give the curve of figure 3, but experimentally the graph plotted is of ρ^2 against $\mathrm{cosec}^2\beta s$, and this shows a crossing of the two branches at some point, since, at any particular value of ρ^2, the value of βs is too small.

Figure 4 is a graph of the simplified function

$$\rho^2 = \mathrm{cosec}^2(\theta+10°) + K\,\cot\,(\theta+10°)$$

plotted against $\mathrm{cosec}^2\theta$ and should be compared with figure 3.

The point of crossing over depends upon the value of b and comes nearer to the bend of the curve as b increases. The width of the loop, however, that is the value of K, also affects the position of this point, so that it does not seem possible, as was hoped at first, to get any information about a or b from the form of this curve.

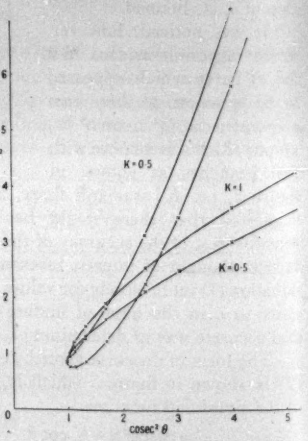


Figure 4.

Considering now the case a finite, $b=0$, which gives

$$\rho^2(1+\sinh^2 a\,\mathrm{cosec}^2\,\beta s) = K_1 + (K_1\sinh^2 a + K_2)\,\mathrm{cosec}^2\,\beta s + 2K_3\,\cot\,\beta s,$$

the general effect of the finite value of a can be seen.

Owing to the multiplying factor on the left-hand side which is always greater than unity, the ordinates will be reduced by a fraction which increases with $\mathrm{cosec}^2\,\beta s$ so that there will be a downward curvature superimposed upon both branches of the curve. At the same time there will be a change in the slope of the whole graph due to the change in the constant which multiplies $\mathrm{cosec}^2\,\beta s$ on the right-hand side.

Figure 5 shows the general curve, plotted for arbitrary values of the constants with both a and b finite. Figure 6 is one of the experimental curves for comparison.

It was concluded then that the copper plate was not a sufficiently good short circuit, and an estimate of the values of the constants a and b was made from the nearest approach to a straight line. These values were not large.

A copper plate of larger size was tried and gave a slight improvement, but evidently the attachment of the plate to the lines was contributing to the reflection coefficient and this quantity could not be reduced to zero in our experiments.

In measuring the reflection coefficient of an unknown impedance Z the quantity b is obtained from the displacement of the minimum of the (ρ^2, s) curve when the short circuit is replaced by Z. To find a a straight line plot $1/(\rho^2-K_1)$

against $\mathrm{cosec}^2 \beta s$ is required and this can only be obtained at the critical separation. Hence, if the lines cannot be satisfactorily short circuited, the critical separation is not accurately obtainable, and there is a certain amount of error in the determination of the reflection coefficient. The error is likely to be greater in a than in b, since the latter quantity is found more directly. Unfortunately it is just the quantity a which is the important one in the case of those dielectrics which show absorption.

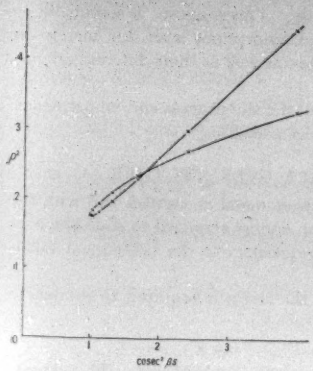


Figure 5.

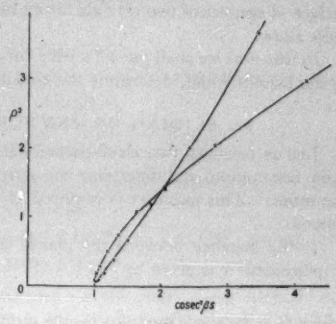


Figure 6.

The experimental work to test these points has been carried out by Miss E. B. Harriss and forms part of the thesis submitted by her for the M.Sc. degree of the University of London. The author wishes to thank her for her help, and also to thank Professor Flint, with whom the problems have been discussed.

REFERENCE

WILLIAMS, G., 1944, *Proc. Phys. Soc.*, **56**, 63.

Slip at Grain Boundaries and Grain Growth in Metals

By N. F. MOTT,
H. H. Wills Physical Laboratory, Bristol

MS. received 6 *January* 1948

§ 1. INTRODUCTION

In a recent paper from the Chicago Institute for the Study of Metals, Ting-Sui-Ke (1947) gives an expression for the rate of slip at the grain boundaries in pure polycrystalline aluminium; this is derived from measurements of the internal friction and some other anelastic phenomena. The expression obtained for the velocity of slip at each boundary is

$$v = A\sigma e^{-B/RT} \text{ cm/sec.}, \qquad \ldots\ldots (1)$$

where σ is the stress applied at the boundary in dynes/cm^2, $B = 34{\cdot}5$ kcal/gm.atom, and $A \simeq 18$ (cm/sec.)/(dyne/cm.).

Ke also points out the following fact : let it be assumed for purposes of calculation that between the grains there is a layer of fluid of thickness t and viscosity η. Then the velocity v of flow should be given by the formula

$$v = \sigma t/\eta. \qquad \ldots\ldots(2)$$

Suppose then that we extrapolate the formula (1) deduced from experiment to the melting point of aluminium, and then compare the results with (2), taking for η the viscosity of liquid aluminium at the melting point. We may deduce a value for t. This comes out to be a few angströms only. This suggests, as is probable on other grounds, that the boundary is not a thick amorphous layer, but merely the surface of contact of two crystals fitting together as well as their different orientations allow.

In this note we shall put forward a theory of the slip process and, in particular, of the factors which determine the constants A and B.

§ 2. SLIDING OF ONE SURFACE OVER ANOTHER

Let us consider two close-packed surfaces of metal in contact and with the same orientation, and determine the activation energy required to slide one over the other. This quantity is required as a preliminary to the subsequent calculations.

If the distance between the planes is h, the stress σ required to produce a displacement x is given by

$$\sigma = Gx/h,$$

where G is the shear modulus for the direction of shear considered. This expression is valid only if x is small. For large x we may set as an approximation

$$\sigma = (Ga/2\pi h)\sin(2\pi x/a),$$

where a is the distance to the nearest stable position of equilibrium; the use of the sine function gives an upper limit to σ. Thus the activation energy per unit area is, to this approximation,

$$\int_0^{\frac{1}{2}a} \sigma\, dx = Ga^2/2\pi^2 h.$$

If $2d$ is the lattice constant, then we may set

$$d = a, \qquad a = 2d/\sqrt{3},$$

and the number $(1/\omega)$ of atoms per unit area is $1/\sqrt{3}\,d^2$. Therefore the activation energy nU required to slide n atoms of the top plane over the lower plane is

$$nU = 3nGa^3/4\pi^2.$$

For aluminium $a = 2{\cdot}02 \times 10^{-8}$ cm., and since the metal is isotropic we may set for G the shear modulus of the bulk material :

$$G = 2{\cdot}5 \times 10^{11} \text{ dyne/cm.}$$

This gives $\qquad U = 0{\cdot}066 \text{ ev.} = c.\ 1{\cdot}5 \text{ kcal.}$

The smallness of this value is to be noted; it is much smaller than the heat of vaporization (76 kcal.), and of the order of the heat of fusion (2·55 kcal.).

§ 3. THEORY OF INTERCRYSTALLINE SLIP

If two crystal planes are in contact, but cannot fit owing to different indices or orientation, one may suppose that the surface of contact is divided into islands

where the fit is reasonably good, separated by lines near which fit is bad. We may suppose that on one of these islands the top plane can move relatively to the one underneath, and that this is the elementary act responsible for intercrystalline slip. With this hypothesis the rate of slip might be estimated as follows :

Let each island contain n atoms and thus cover an area $n\omega$. In the absence of stress the activation energy for slip is nU. In the presence of a stress σ it is $nU \pm \frac{1}{2}\sigma an\omega$, it being supposed that $\frac{1}{2}a$ is the distance to the potential barrier separating one position of equilibrium from the next. Thus the number of times per second that each island moves a distance a in the direction of the stress is

$$\nu \exp\{-(nU - \sigma na\omega)/\boldsymbol{k}T\},$$

and in the opposite direction

$$\nu \exp\{-(nU + \sigma na\omega)/\boldsymbol{k}T\},$$

where ν is the frequency of atomic vibration. It follows that the rate of slip is

$$\nu e^{-nU/\boldsymbol{k}T} \sinh(\sigma na\omega/\boldsymbol{k}T).$$

If σ is not too great this becomes

$$(\nu a^2 n\omega\sigma/\boldsymbol{k}T)e^{-nU/\boldsymbol{k}T}.$$

If nU is to be identified with the observed 34·5 kcal., we must have

$$n = 34{\cdot}5/1{\cdot}5 \simeq 23.$$

With $\nu \simeq 10^{12}\,\text{sec}^{-1}$, $a \simeq 2 \times 10^{-8}$ cm., $\omega \simeq 7 \times 10^{-16}\,\text{cm}^2$, $T \simeq 600°$, we get for A

$$A = \nu a^2 n\omega/\boldsymbol{k}T \simeq 6 \times 10^{-5}.$$

In this treatment it has been assumed that σ at the point of slip is equal to the applied stress. Some stress magnification is possible, in which case A would be somewhat greater.

The objections to this treatment are :

(*a*) The calculated value of A is too small by a factor 10^4–10^5.

(*b*) The observed fact that at the melting point the slip is the same as that which would be given by a monomolecular layer of liquid appears in the theory as an *accident*. The mechanism of flow in liquid aluminium cannot be anything like that sketched here, because the temperature dependence is 10–20 times smaller; the viscosity of liquid metals depends on temperature according to the formula

$$\sigma = \sigma_0 e^{W/\boldsymbol{k}T},$$

where W is of the order of the latent heat of fusion (Frenkel 1946).

We must therefore modify our hypothesis. Let us suppose that the elementary act which allows slip to occur is the *disordering* of atoms round each island where fit is good. The free energy F necessary to do this will approach zero at the melting point and nL at the absolute zero of temperature; here L is the latent heat of fusion per atom. At any other temperature, let us assume F to be given by

$$F = nL(1 - T/T_M),$$

where T_M is the temperature of melting. Let us assume also, since the disordering will result in a slip through a distance a, that a stress σ will decrease or increase F by $\pm \frac{1}{2}\sigma n\omega a$. Then the rate of slip is now

$$v = 2\nu a \exp\{-nL(1 - T/T_M)/\boldsymbol{k}T\} \sinh(\sigma n\omega a/2\boldsymbol{k}T),$$

which for small σ reduces to

$$v = \frac{\nu a^2 n\omega\sigma}{\boldsymbol{k}T} \exp\left(\frac{nL}{\boldsymbol{k}T_M}\right) \exp\left(\frac{-nL}{\boldsymbol{k}T}\right).$$

For aluminium the latent heat L is 2·55 kcal/mol, so that n (the number of atoms to be disordered) should be about 14.

The factor $\exp(nL/\boldsymbol{k}T_M)$ is $\exp(34\,000/2\times 930)\simeq 10^8$. One is thus able to explain the large factor outside the exponential. One also sees why the surface layer behaves like a layer of liquid at the melting point, since the activation energy for melting small volumes there disappears.

§ 4. VELOCITY OF RECRYSTALLIZATION

Anderson and Mehl (1945) have shown that the velocity of growth of new crystals during recrystallization of cold-rolled aluminium is given by

$$v = v_0 e^{-Q/RT},$$

where $Q\simeq 55$ kcal. and $v_0\simeq 10^{14}$–10^{16} cm/sec. Burgers and the present author (1947) have suggested that this very high value of v_0 indicates that some process which depends on temperature through the factor $e^{-Q/RT}$ triggers off the change of crystal form of a whole mosaic block, which then changes its crystal structure to that of the growing crystal without the intervention of any other thermally activated process.

An alternative and perhaps more plausible explanation can be given along the lines of the last section. Let us assume that the crystal grows at each step by the disordering or melting of a group of n atoms of the cold-worked material which then takes up the structure of the growing crystal. We have at present no theoretical means of determining n; as in the previous problem it must be the value for which the free energy of disordering, including surface energy, is a minimum.

We should thus expect as in the last section a rate of crystal growth

$$va\exp(Q/RT_M)\exp(-Q/RT). \qquad \ldots\ldots(3)$$

With $va\simeq 10^{+3}$ cm/sec., and the first exponential factor equal to 10^{13}, this gives $10^{16}e^{-Q/RT}$, which is of the observed order of magnitude.

This analysis assumes that the local melting always results in an addition to the growing crystal and never the reverse process. To test this assumption, let q be the mean energy of cold work per atom. The condition for the validity of the assumption is then $nq\gg\boldsymbol{k}T$. Otherwise (3) should be multiplied by $(1-e^{nq/\boldsymbol{k}T})$. For heavily cold-worked aluminium according to Taylor and Quinney (1934) q is 30 cals/mol. With $n\simeq 20$ this gives $nq/\boldsymbol{k}T$ at 600° K. of the order $\frac{1}{2}$. For lightly rolled aluminium the value should be considerably less, but the order of magnitude of (3) will not be affected.

REFERENCES

Anderson, W. A., and Mehl, R. F., 1945, *Amer. Inst. Min. and Met. Engng.*, Tech. Pub. No. 1805.

Barrett, C. S., 1943, *Structure of Metals.*

Burgers, W. G., 1947, *K. Ned. Ak. Wet.*, **50**, 452.

Frenkel, J., 1946, *Theory of Liquids* (Oxford: Clarendon Press).

Taylor, G. I., and Quinney, H., 1934, *Proc. Roy. Soc.* A, **143**, 307.

Ting-Sui-Ke, 1947, *Phys. Rev.*, **71**, 533.

LETTERS TO THE EDITOR

Paramagnetic Resonance at Low Temperatures in Chromic Alum

The use of centimetre wavelengths to determine the small splitting (due to the crystalline electric field) of the basic spin quadruplet of Cr^{+++} in potassium chromic alum has been described by Bagguley and Griffiths (1947), who find a value of 0·12 cm^{-1} for the separation between the $\pm 1/2$ and $\pm 3/2$ levels. The same authors (unpublished) have obtained the value of $0{\cdot}13_5$ cm^{-1} for the analogous splitting in ammonium chromic alum; independently Weiss, Whitmer, Torrey and Jen-Sen Hsiang (1947) give a value $0{\cdot}15 \pm 0{\cdot}01$ cm^{-1} for this salt, obtained by a similar method. These values differ considerably from those obtained at low temperatures by other methods: viz. $0{\cdot}16-0{\cdot}19$ cm^{-1} for potassium chromic alum (Bleaney 1939, Du Pré 1940, Casimir, de Haas and de Klerk 1939, Starr 1941, Broer 1947, Gorter *et al.* 1940, 1942) and 0·34 cm^{-1} for ammonium chromic alum (Starr 1941). The application of centimetre wavelengths to the determination of the change of splitting with temperature was therefcre undertaken by the authors.

A fixed tuned cavity is used, resonant in the H_1 mode at about 3 cm. wavelength; it is excited by probes at the ends of small concentric line feeders. The cavity is surrounded by a dewar vessel, containing liquid oxygen, liquid hydrogen etc.; its temperature is determined by a vapour pressure thermometer, or by a thermocouple calibrated at various fixed points.

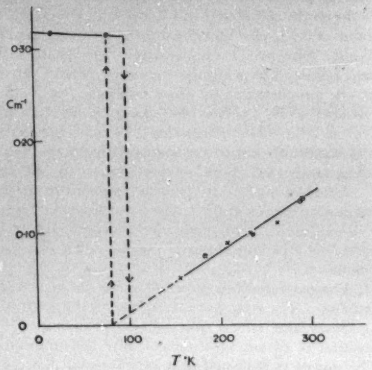
The magnification factor Q of the empty resonator is about 13 000 at 90° K. and 15 000 at 20° K. The magnetic absorption in the crystal rises with $1/T$, and it is therefore possible to obtain large effects from very small crystals. The crystal is placed in the resonator in a position of maximum magnetic field and minimum electric field (to avoid dielectric loss). The power transmitted through the cavity at resonance is observed by means of a silicon-tungsten rectifier while a variable magnetic field is applied, perpendicular to the R.F. magnetic field.

As the temperature is lowered, the absorption curve for ammonium chromic alum remains similar to that observed at room temperature, except that the side peaks due to the Stark splitting move steadily inwards towards the main peak, showing that the splitting is decreasing. At about 80° K., however, a transition sets in, which requires several minutes for its completion. It is accompanied by a considerable change in the resonant frequency of the cavity, and the absorption curve obtained on applying a magnetic field after the transition is complete is quite different from that obtained at higher temperatures. On warming up slowly, the reverse transition does not occur until about 100° K., after which the magnetic absorption curve reverts to that previously obtained at this temperature. These transitions are repeatable, but the exact transition temperature varies by a few degrees; in general the hysteresis between the transition temperatures for cooling and warming decreases after the operation has been repeated a number of times.

There is little doubt that this effect is the same as that found by Guillien (1939), who discovered a discontinuity in the dielectric constant, accompanied by jumps in the dielectric loss and specific heat, showing temperature hysteresis. The change in resonant frequency of the cavity observed by us is consistent in sign with this dielectric constant change, though comparatively small since the crystal projects only slightly into the R.F. electric field. The effects are attributed by Guillien to " co-operative relaxation of librational restraint on water or water-containing complexes with rising temperature ". If the water molecules involved are those forming the octahedron around the Cr^{+++} ion, the great influence of this transition on the Stark splitting of this ion can be understood, since Van Vleck (1939) has shown that the splitting can be explained by the slight distortion of the octahedron, together with other contributions due to the Jahn-Teller effect and to more distant ions. The transition has also been observed by Kraus and Nutting (1941), who found that the crystal often became practically opaque, as though shattered into an enormous number of tiny crystals; this is consistent with our observations. This shattering process does not, however, alter the magnetic absorption spectrum observed at higher temperatures.

Below the transition point, the magnetic absorption curve observed with the external magnetic field along the normals to the (111), (110), (100) crystal planes is quite inconsistent with that to be expected from the theory of Broer (1942), assuming a small trigonal field superposed on the cubic field. Since the results at room temperature are in good agreement with this theory, one must assume that at low temperatures the crystalline electric field is differently aligned or has different symmetry. The quantitative interpretation of the magnetic absorption curve is therefore not possible at this stage. In this salt, however, there is considerable absorption at 3 cm. wavelength in zero magnetic field below the transition point, showing that the splitting is of this order. By examining the absorption curve in low fields, the splitting has been determined as $0{\cdot}314\ cm^{-1}$ just below the transition point, increasing slightly to $0{\cdot}317\ cm^{-1}$ at 20° K. ; the probable error in these values is less than 1%. Since, by Kramers' theorem, an electric field cannot do more than decompose the spin quadruplet into two doublets, these values for the position of maximum absorption in zero field must correspond to the Stark splitting, irrespective of the nature of the crystalline field.

The variation of the Stark splitting down to 20° K. is shown in the figure for ammonium chromic alum. Above the transition point the splitting varies linearly with temperature, and is, presumably, associated with rotation of the water molecules. Below the transition point, the slight change of the splitting may be due to thermal contraction, and it seems unlikely that there will be any substantial change below 20° K. The results of preliminary measurements on potassium chromic alum show that there is a similar decrease in the splitting on cooling downwards from room temperature, but no sudden transition occurs. Instead, abnormally weak side peaks appear in the absorption curve which, though not consistent with a trigonal field directed along the space (body) diagonal of unit cell, appear to correspond to a splitting of the same order as that obtained by other methods at low temperatures. Measurements are in progress at longer wavelengths to determine this splitting directly by a method similar to that used for the ammonium alum.



Variation of Stark splitting in ammonium chrome alum with temperature.

⊡ Bagguley and Griffiths.
× Before shattering.
⊙ After shattering.

B. Bleaney.
R. P. Penrose.

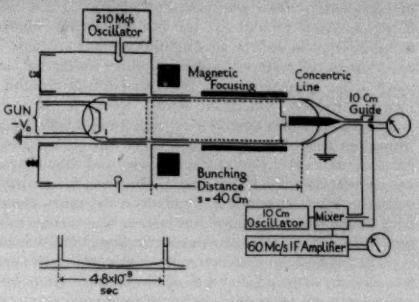
The Clarendon Laboratory,
Oxford.
19 December 1947.

Bagguley, D. M. S., and Griffiths, J. H. E., 1947, *Nature, Lond.*, **160**, 532.
Bleaney, B., 1939, unpublished, quoted by Van Vleck, 1940.
Broer, L. J. F., 1942, *Physica*, **9**, 547 ; 1947, *Ibid.*, **13**, 353.
Casimir, H. B. G., de Haas, W. J., and de Klerk, D., 1939, *Physica*, **6**, 365.
Dijkstra, L. J., Gorter, C. J., and Van Paemel, O., 1942, *Physica*, **9**, 673.
Du Pré, F. K., 1940, *Physica*, **7**, 79.
Gorter, C. J., Dijkstra, L. J., and Groendijk, H., 1940, *Physica*, **7**, 625.
Guillien, R., 1939, *C.R. Acad. Sci., Paris*, **209**, 21.
Kraus, D. L., and Nutting, G. C., 1941, *J. Chem. Phys.*, **9**, 133.
Starr, C., 1941, *Phys. Rev.*, **60**, 241.
Van Vleck, J. H., 1939, *J. Chem. Phys.*, **7**, 61 ; 1940, *Phys. Rev.*, **57**, 426.
Weiss, P. R., Whitmer, C. A., Torrey, H. C., and Jen-Sen Hsiang, 1947, *Phys. Rev.*, **72**, 975.

Milli-microsecond Pulse Generation by Electron Bunching

Work is in progress in this laboratory on a project for coincidence counting of nuclear particles with a resolution of the order of 10^{-9} seconds. For a study of transients of this duration, the first requirement is a technique of milli-microsecond pulse generation; these periods of time are below the limit of normal electrical pulse production methods. Equipment has been built for the production of these pulses by bunching electrons with alternating voltages of frequency 210 Mc/s., on the assumption that the mathematical theory of electron bunching applies (Webster 1939, Harrison 1947).

The figure shows in diagram the electronic arrangement, which is in effect a klystron with the resonant catcher replaced by a non-resonant collector in the form of a coaxial line. Power from a 210-Mc/s. oscillator is fed into a coaxial line resonator short-circuited at one end and capacitively terminated at the other by two grids 1 cm. apart, across which R.F. potentials up to 400 volts can be developed. An electron beam, after modulation by these grids, is magnetically focused on to the end of a 70-ohm coaxial line a distance $s=40$ cm. away. The outer conductor is terminated in a grid through which the beam passes before reaching the inner conductor 4 mm. away.



Bunching pulse generator and associated circuits in block diagram form.
Inset: Theoretical pulse

The bunched beam passing across the capacity between the beam and the conductor induces a voltage pulse which is then propagated down the line. The line is terminated in a probe, which feeds a length of 10-cm. waveguide through a "door knob" transformer, and a milliammeter by which the D.C. beam current, of the order of 2 ma., is read. The waveguide is terminated by a 10-cm. mixer with local oscillator, and a 60-Mc/s. I.F. amplifier with diode current meter (Collie, Hasted and Ritson 1948). This serves as a narrow band power measuring device with which it is possible to pick out a harmonic component of the pulse. The peak alternating voltage in the bunching pulse generator, that is the peak energy increments given to the electrons, can be measured directly by back-biasing the electrode on which the beam impinges until no current is observed either with or without the alternating voltage at the buncher grids.

The theory of electron bunching predicts that the nth harmonic of the alternating component of the current is proportional to $J_n(nx)$; x is the bunching parameter, given by the relation $x=V_1 s\pi f/V_0 v_0$ where v_0 is the velocity of the electrons accelerated to a potential V_0, and V_1 is the alternating voltage of frequency f. With the local oscillator tuned to a wavelength of 10 cm., maximum power in the 14th harmonic (of the order of 10^{-4} watts) was obtained with values of $V_1=250$–270 v. at $V_0=2000$ v., and $V_1=140$ v. at $V_0=1350$ v.; both of these results correspond to $x=1{\cdot}3$, $nx=16{\cdot}8$. The first maximum of the 14th Bessel function occurs at $nx=16$, so that the agreement with theory is satisfactory. The theoretical pulse, the shape of which it is hoped to verify with viewing equipment, is shown in the inset to figure 1 to be of the order of 2×10^{-10} seconds long. Responses have been plotted for varying values of nx, and lie on a curve of roughly the same form as that of the Bessel function, although observation of maxima beyond the first has not yet been attempted.

From the power observed in the 14th harmonic, and the agreement with theory, we should expect that this device will be capable of supplying pulses of variable width and of duration short enough for testing wide-band coincidence counting circuits.

The Clarendon Laboratory, J. B. HASTED.
Oxford.
11 December 1947.

COLLIE, C. H., HASTED, J. B., and RITSON, D. M., 1948, *Phys. Soc. Proc.*, **60**, 71.
HARRISON, A. E., 1947, *Klystron Tubes* (London and New York: McGraw-Hill).
WEBSTER, P. L., 1939, *J. Appl. Phys.*, **10**, 501.

REVIEWS OF BOOKS

Radar System Engineering, edited by LOUIS N. RIDENOUR. Pp. xviii + 748 with 646 figures. (New York: McGraw Hill Book Co. Inc., 1947.) 37*s*. 6*d*. net.

During World War II a tremendous research and development effort went into the development of radar and related techniques, and resulted not only in hundreds of radar sets being available for military use but also in the collection of much valuable data and new techniques in the electronics and high-frequency fields. This effort was a splendid example of co-operative research under war-time conditions, for each new advance was rarely the work of an isolated worker, but generally the combined efforts of a very large group of people: British, American, Army, Navy, R.A.F., civilian, Government, and industry, working in different laboratories, and often at widely separated locations. A very large and important part of this combined effort was undertaken by the Radiation Laboratory of the Massachusetts Institute of Technology, and they have now played a further important part by undertaking the heavy responsibility of preparing for publication the results of this collective work. As originally planned the series comprised twenty-three volumes dealing with theory and techniques, but later it was decided to extend the series by the addition of a further five books dealing with radar and applied systems.

The present book, which is the first of the series to be published, is also the first of the five dealing with radar and applied systems. It is intended to serve as a general treatise and reference book on the design of radar systems: it deals, in fact, with the principles and practice of radar. Much of the work has been written by Professor Ridenour himself, but he has been assisted in this monumental task by many other writers, thirty-two of whom have contributed to this first volume.

The aim of the book, as stated in the preface, is to cover the collective efforts of the many workers, both in the United States and in Britain, Canada, Australia and other British Dominions, but in spite of this, the book tends to be rather overweighted with American applications of the new techniques, and this may detract from its value to the British reader. Furthermore, although various references to the original literature are given throughout the book, the references mentioned are usually to obscure reports (e.g. M.I.T. Radiation Laboratory Reports) not generally available in this country. The book would have been of more value to the British reader if reference could have been made to the excellent literature on the subject published last year by our Institution of Electrical Engineers. However, these are small points, and it can be said at once that this is a very good book, which should be of real value to all those wanting to make themselves conversant with the principles and practice of the subject.

Chapter I is a general introduction and explains clearly and concisely "what radar does" and "how it works". It then goes on to give details of the components of a radar system, the types of radar developed during World War II and the performance obtained. The author departs a little from the usual custom by referring at the outset in his explanation to a microwave pulse radar system, rather than to the simpler and historically earlier metric-wave systems. This insistence on the importance of microwaves from the start is quite reasonable. Considerably more work has gone into pulse radar than into any other kind, and by far the greater part of this work has been concerned with microwave pulse radar. Moreover, it is now recognized that for almost all radar purposes microwaves offer advantages. This first chapter is very well written and the explanations are good, but one small mis-statement has been made. On page 11, in referring to super-refraction conditions, the author states "that such phenomena are relatively rare and essentially unpredictable". This statement would be true in referring to this country; it is probably also true in referring to conditions in North America, but it is obviously not true in relation to conditions in more tropical countries. The Royal Air Force have observed super-refraction conditions repeated for several weeks in the year, for several years in succession, in certain parts of India, and so for certain parts of the world, at least, the increase in range of a radar set beyond the optical horizon is a predictable phenomenon.

Chapter II contains a treatment of the radar equation according to which the range (R) of a radar set is a function of the transmitted power (P), the gain of the antenna (G), the wavelength (λ), the radar cross-section of the target (σ) and the minimum detectable received signal (S). The author writes down his formula as

$$S=\left(\frac{PG}{4\pi R^2}\right)\left(\frac{\sigma}{4\pi R^2}\right)\left(\frac{G\lambda^2}{4\pi}\right), \qquad \ldots\ldots \text{(A)}$$

and proceeds to discuss in a very thorough fashion the consequences of the equation, and the choice of the various parameters for particular purposes. He discusses beams of special shapes, the limitations of the receiver, receiver noise and the statistical problem, and the modification of the range equation called for when " free-space " propagation conditions are not applicable. He discusses one case in detail, namely the calculation of the maximum radar range of an aerial target when the propagating path is over a flat, perfectly reflecting surface, and shows that in this case the " free-space " formula becomes

$$S=P\frac{G^2\lambda^2\sigma}{(4\pi)^3R^4}\cdot 16\sin^4\left(\frac{2\pi h_1h_2}{\lambda R}\right), \qquad \ldots\ldots \text{(B)}$$

and he discusses, in detail, how accurately the flat, perfectly reflecting conditions laid down in his analysis are fulfilled in practice, and therefore how reasonable it is to use the new equation. The treatment, so far as aerial targets are concerned, is thorough and excellent, but the author fails to point out the limitation of his analysis to ship targets, and except for a casual reference to figure 2.12, the reader is left with the impression that equation (B) may be used for ship targets. This is very unsatisfactory. The reviewer would have liked to have seen the important case of the ship target considered in detail with proper comment on the special conditions which arise in this case (namely that the target is not placed in a constant incident field, but in one which varies across the surface of the target), and the modifications to the concept of target cross-section made necessary. A satisfactory treatment of this problem was published recently by Wilkes and Ramsay (*Proc. Camb. Phil. Soc.*, 1946, p. 220).

Chapter III gives a detailed treatment of the properties of radar targets, and after a consideration of targets of simple shapes, e.g. sphere, corner reflector etc., some discussion is included on methods of target shaping and the use of absorbent materials to diminish cross-section. The chapter then goes on to discuss the radar returns from more complicated targets such as aircraft, " window ", rain, land, and the sea, and this is illustrated with many excellent photographs of the radar returns actually obtained. The use of radar " land-painting " is considered very fully in connection with the important applications of this technique to navigation problems.

Chapter IV discusses the limitations of pulse radar, and Chapter V gives a detailed treatment of the many C.W. radar systems, which includes many critical inter-comparisons of the performance of C.W. and pulse systems. The C.W. systems considered include simple Doppler systems, range-measuring Doppler systems, frequency-modulation range measuring systems, multiple-target frequency-modulation systems, pulse-modulated Doppler systems etc. One small mis-statement occurs sufficiently frequently to be annoying, namely that a frequency-modulation system will work down to zero range. There is a minimum range for a F.M. system but it is not zero, but equal to $c\delta f/2F\,\Delta f$, where δf is the beat frequency, F is the modulation frequency and Δf is the frequency swing. This reduces to $c/2\Delta f$ when $\delta f=F$. The reviewer would have liked to have seen this point dealt with properly. The corresponding figure for a pulse system is $\frac{1}{2}c\tau$, where τ is the pulse length. For a normal F.M. system the limitation is usually of the order of 1 metre, whereas for a pulse system the limitation in range might be as great as 100 metres ; hence all the author's statements about the advantages of C.W. systems for the accurate measurement of very short ranges are obviously true.

Chapter VI gives details of the presentation of radar information, and most readers will find this a valuable introduction to the various applications of radar methods. Chapter VII considers the employment of radar data, and gives examples of radar organization, viz. radar in the R.A.F. Fighter Command and the U.S. Tactical Air Commands. It is very pleasing to the reviewer to find this important material included. The technical design of a radar set is relatively straightforward to the specialist at least, but the employment of radar data in the

best possible way is far less well understood, and Professor Ridenour is quite correct when he states that the major improvements to be looked for in the use of radar during the next few years will lie, for the most part, not in technical radar design, but in the field of fitting the entire radar system, including its operational organization, to the detailed needs of the use and the user. This obviously applies to both military and peace-time applications.

Chapter VIII continues with a consideration of radar beacons, and gives a good introduction to the subject. The whole of volume 3 of the series is also devoted to this subject. Chapters IX to XIV take up the leading design considerations for the various important components that make up a radar set, and deal respectively with aerials, magnetrons, R.F. components, radar receivers, indicators and power supplies. The treatment is thorough and written throughout by specialists in their subjects. The reviewer would have liked the needs of the British reader to have been considered, by finding a few more references to British practice, and he would have expected the 70-page chapter on the magnetron developments to have mentioned that the all-important cavity resonator magnetron was invented by Professor Randall, our recent Duddell medallist, but bearing in mind the emphasis on co-operative efforts, perhaps this is not so unfair. Chapter XV is one of the most valuable in the book and gives some examples of radar system design practice. The authors (R. G. Heib and R. L. Sinsheimer) consider the complete design problem from a consideration of the formulation of requirements, to the choice of the various radar parameters, and to the detailed problems of the design of the various component parts of the equipment. This chapter has been carefully written and will repay careful study. Finally the book closes with a consideration of two new and important ancillary techniques—moving target indication and the transmission of radar displays to a remote indication by radio means.

The book is well produced and illustrated and can be recommended to any serious student of radio. It is hoped that some of the criticisms made above will be dealt with in future editions. DENIS TAYLOR.

Radioactivity and Nuclear Physics, by J. M. Cork. Pp. x+313. (New York: Van Nostrand Co. Inc., 1947.)

This book will be easily read—and with a fair measure of assent—by anyone familiar with the subject, yet it is a tantalizing and a disappointing book. It lacks logical form: there is a great deal that passes for information, very little which seeks to be educative. The serious student will not find that the book answers his questions, though it will certainly give rise to many, in a bewildering sequence. All the fashionable terms are freely used—curium is mentioned on p. 3, fission on p. 6, the relativistic change of mass with velocity on p. 14, and the meson and the neutrino on the following page. But that surely is the wrong place for them, unless a very casual acquaintance with words is all that the reader requires.

"It has been the aim of this text to offer an authentic historical development of each topic": thus the preface. This much must be granted, that the aim is evident throughout; nor is its success negligible (most obvious perhaps in Chapter 3 which describes the evolution of large-scale apparatus for particle acceleration), but the historian is at fault, sometimes, and at others is uncritical. It would seem strange that the earliest reference to Rutherford's work in radioactivity is of date 1902, and the almost literal statement that the experiment of Ellis and Wooster (1927) was planned in order to test the neutrino hypothesis is a peculiar reversal of history. Yet the reader should be grateful that the historical approach is followed consistently enough, whenever occasion offers, and the copious references which appear as footnotes to the text are a good feature of the book.

At the end appendices give tables of isotopes (stable and unstable), of fundamental constants, of numerical data concerning electrons, protons and α-particles over a considerable range of energy, and of the characteristic X-ray excitation energies of the elements titanium to uranium. These tables also are good (incidentally the table of isotopes includes some data regarding recently discovered isotopes of uranium and the transuranics which, so far as the present reviewer is aware, have not hitherto been published either in this country or in America). But why exclude rhenium and bismuth from the table of X-ray excitation energies? This is rather a frivolous question, but it should set the user of the tables on his guard—for there are, perhaps inevitably, some errors as well as omissions.

From what has been said in praise of the industry which has gone to the compilation of lists of references and useful tables of data it will be evident that it is the greater pity that

the general level of the text is not better. But criticism cannot be avoided: the text too frequently appeals to the memory of the student, not to his understanding. And even his memory is not encouraged by a repetitiveness which so often introduces inconsistency: the maximum energy of the β-particles of radium E is given as 1050 kev. on p. 124 and as 1·17 Mev. on p. 133, the discovery of the positron is attributed to Anderson in 1933 on p. 153—and in almost identical language the date is given as 1932 on p. 124 (incidentally, on the former page the relativistic theory of the electron is dated 1934—six years too late), the energy release in the reaction $^{1}{}_{5}^{0}\mathrm{B}\ (n,\ \alpha)\ {}_{3}^{7}\mathrm{Li}$ is quoted as 2·5 Mev. on p. 177 and as 3·0 Mev. on p. 181, that for the reaction ${}_{1}^{3}\mathrm{H}\ (d,\ n)\ {}_{2}^{3}\mathrm{He}$ as 3·17 Mev. on p. 178 and 3·16 Mev. on p. 198 (the "best" value is in fact 3·23 Mev.), that for ${}_{4}^{9}\mathrm{Be}\ (d,\ n)\ {}_{5}^{10}\mathrm{B}$ as 4·1 Mev. on p. 178 and 3·74 Mev. on p. 198, and on p. 210 it is stated that the specific ionization due to an electron is independent of its energy when the energy is high, whereas on p. 229 we have the correct statement that the ionization increases with increasing energy in this range. These are pure inconsistencies, which diminish confidence when they are detected; of the outright errors, which the unwary student might not find it so easy to detect, it will suffice to quote one in full (p. 125)—"The upper limit of the spectrum for the ^{64}Cu positrons is 0·66 Mev. compared with 0·58 Mev. for the negatives. The difference may be due to the action of the positive nucleus on the departing particles, being an attraction for the negatron and a repulsion of the positron". It must be admitted that there is a superficial air of verisimilitude in this wholly misleading statement.

N. FEATHER.

Tables of Spherical Bessel Functions. Vol. II. Prepared by the Mathematical Tables Project of the National Bureau of Standards. Pp. xx+328. (New York: Columbia University Press, 1947.) $7.50.

The tabulation of the functions $\sqrt{(\pi/2x)}\,.\,J_{\nu}(x)$ is continued*, and this volume covers the ranges

$$\pm 2\nu = 29\ (2)\ 43; \qquad x = 0\ (0{\cdot}01)\ 10\ (0{\cdot}1)\ 25;$$
$$\pm 2\nu = 45\ (2)\ 61; \qquad x = 10\ (0{\cdot}1)\ 25.$$

Apart from a few entries near the zeros, the values are given to eight or more significant figures for $x<10$, and to seven for $x>10$. Second differences, modified second differences, or second and fourth differences, are given to facilitate interpolation, except in ranges from $x=0$ where the functions vary very rapidly. To enable interpolation in these latter ranges (and also in the ν direction) to be performed, values of

$$\Lambda_{\nu}(x) = 2^{\nu+\frac{1}{2}}\,\pi^{-\frac{1}{2}}\,\Gamma(\nu+1)\,J_{\nu}(x)$$

are given for

$$2\nu = 1\ (1)\ 41\ (2)\ 61; \qquad x = 0\ (0{\cdot}1)\ 10;$$
$$2\nu = 1\ (2)\ 61; \qquad x = 10\ (0{\cdot}1)\ 25;$$
$$-2\nu = 29\ (2)\ 33; \qquad x = 0\ (0{\cdot}1)\ 9{\cdot}5\ (0{\cdot}05)\ 10\ (0{\cdot}1)\ 25;$$
$$-2\nu = 35\ (2)\ 61; \qquad x = 0\ (0{\cdot}1)\ 25;$$

the first group to nine decimal places, the second to (mostly) seven significant figures.

Finally the zeros of $J_{\nu}(x)$ and $J_{\nu}'(x)$ obtained by inverse interpolation from the main tables of this and the preceding volume, and the value of each function at the zero of the other, are given to six (occasionally more) decimal places.

The functions here tabulated are important in, for instance, solutions of the wave equation in spherical polar coordinates, and are to be welcomed as extending the range of the tables already given in volume I.

W. G. B.

Table of the Bessel Functions $J_0(z)$ and $J_1(z)$ for Complex Arguments. Second edition. Prepared by the Mathematical Tables Project of the National Bureau of Standards. Pp. xliv+403. (New York: Columbia University Press, 1947.) $5.

This is substantially a reprint of the first edition, but some errors of lettering in the diagrams have been corrected and explicit relations between the functions tabulated along the 45° ray and the *ber* and *bei* functions are now included, with other minor revisions, in the introduction.

W. G. B.

* cf. *Proc. Phys. Soc.*, 1947, **59**, 259.

Time and Thermodynamics, by A. R. UBBELOHDE. Pp. vi+110. (Oxford: University Press, 1947.) 6*s*. net.

The main theme of this well-produced little book is that, in the metrical study of the external world, vision (i.e. the use of any of the so-called ethereal vibrations) is of prime importance, and since vision depends on the absence of radiation equilibrium (in a constant temperature enclosure the world would appear as an undifferentiated uniformity) the conditions of such absence (e.g. the continual increase of entropy) are ultimately the conditions of obtaining metrical knowledge of the universe. " And since this is so it may be anticipated that the trend of events, or direction in time, which is indicated by thermodynamics has its foundations primarily in the thermodynamics of radiation " (p. 77).

The book is disappointing, both in its main thesis and in the development of it. There are other conditions, besides absence of radiation equilibrium, of obtaining metrical knowledge ; for example, the temperature of the body must not exceed, say, 110° F., the atmosphere must not be deprived of oxygen, and so on. It is not clear why, among these conditions, the absence of radiation equilibrium is " prime ". When we come to time (which, by the way, is very much the junior partner, notwithstanding the order of names in the title), the book is even less satisfactory. It appears to be held that entropy enables us to measure spontaneous change. " Historically, the way in which the concept of entropy first aroused an interest outside its immediate technical applications was by providing a quantitative measure for the trend of spontaneous happenings in the particular instance of the direction of spontaneous heat flow from one body to another " (p. 78). This, of course, is nonsense. The second law of thermodynamics merely requires that in an irreversible change from one equilibrium state to another the entropy must increase. It says nothing at all about the *rate* of increase, and therefore provides no quantitative measure of the trend.

Still more unexpectedly, the book abounds in gross errors. Perhaps the most surprising is in the discussion (pp. 83–6) of the expansion of a gas from a volume v_1 to a volume v_2. It is stated that " the ratio of the probability W_1 that the molecules will all remain in the volume v_1 to the probability W_2 of their occupying uniformly the larger volume v_2 which is made available to them, is given by $W_2/W_1=v_2/v_1$ ", and even the recital of an example of patrons choosing theatre seats, in which the absurdity is equally obvious, has not opened Professor Ubbelohde's eyes to it. Again, " the unfolding of the solar system on Laplace's dynamical theory would be completely reversed by reversing the momenta of all the masses taking part " (p. 79). Does Professor Ubbelohde really think that if the bodies revolved the other way the planets would return to a Sun which would expand back to the original nebula ? Again, infra-red solar radiation " is largely responsible for the sun's heat, and for the comfortable use of glasshouses " (p. 61). Professor Ubbelohde surely knows how small a proportion of the sun's heat comes from infra-red radiation, and how negligible its part is in the greenhouse effect. Once more, the uncertainty principle is stated as $\Delta E \,.\, \Delta t \geqslant h$, " where h is Planck's constant in radiation theory, and ΔE is the minimum uncertainty in E, Δt the minimum uncertainty in t " (p. 75). These increments are, of course, any uncertainties at all. The least value for each is zero and the greatest infinity, and only the product has a minimum value. There are other faults of the same type, of which we need add only the inexcusable transformation by which probability conceptions, first properly introduced (p. 83) as terms in which the laws of thermodynamics can be " reformulated ", automatically become " the statistical *foundations* of the Second Law of Thermodynamics " (pp. 87, 97, 101 ; reviewer's italics).

It is difficult to persuade oneself that this book has really come from so competent and well-informed an author as Professor Ubbelohde, but one must believe the Oxford University Press. He is plainly guilty of contempt of subject. Fortunately a reviewer is in the position of jury and not of judge, and we gladly refrain from passing sentence.

HERBERT DINGLE.

Fundamentals of Photography with Laboratory Experiments, by PAUL E. BOUCHER. 2nd edition. Pp. xii+394+liii. (New York: D. Van Nostrand Co. Inc., 1947.)

The first edition of this work was published in 1940 ; the present edition has been broadened, so the preface tells us, by the inclusion of more material on colour photography

and by chapters on ultra-violet and infra-red, x-ray, and motion-picture photography. The book is intended to supply lecture and laboratory notes for a one-semester course in photography for college students. The major part consists of some nineteen chapters ranging from the historical development of photography to cinematography. There follow a series of twenty-five laboratory exercises which follow the same order as the chapters.

The book appears far too ambitious. It contains too much and as a result in the treatment of specific matters it is superficial, and many careless statements occur. The sketch of the historical development of photography is very unsatisfacotry. Fox Talbot is left out entirely and Herschel is not mentioned. Niepce is mentioned, but his really important early work is overlooked. In the next chapter is a heading—" The Reduction of the Latent Image ". One knows something of the development of the latent image, but we do not usually talk about its reduction. Very many more examples could be cited of careless statements, and it has to be remembered that this work is intended as a college text-book. It does not deal satisfactorily with the fundamentals of photography, while at the same time dealing very superficially with many other matters. It is surely unnecessary to provide a description and a wiring diagram of an x-ray generator when no mention is made of the use of intensifying screens in x-ray work. There is no mention of stereoscopic photography, but quite a lot about certain aspects of ultra-violet and infra-red that might well have been left out.

The general plan of the book is good and most of the information sound and well arranged ; it is a great pity that it should be marred by such imperfections as are detailed above.

H. W. GREENWOOD.

Atomic Energy, by R. R. NIMMO. Pp. 201. (London : The Pilot Press Ltd., 1947.) 9*s*. 6*d*.

" The average citizen cannot be expected to understand clearly how an atomic bomb . . . works " says Professor Smyth in the preface to the official United States booklet on Atomic Energy. On the other hand, the Pilot Press and their editor, Dr. Lovell, believe that the " non-specialist public " is capable of keeping abreast of such topics as atomic energy, and they publish this book by Dr. Nimmo to prove their point. We shall not attempt an opinion on this controversial question here, but simply examine the content of Dr. Nimmo's book. It contains an excellent account of nuclear fission, chain reactions and uranium piles, leading up to the atomic bomb and atomic energy plants. A fairly conservative attitude is indicated towards the future of atomic power plants. Ten thousand tons, or more, of uranium would be needed yearly to provide the power requirements of Great Britain, and there is little up-to-date information available, apparently, regarding the size of uranium deposits. Generally speaking, all the information on atomic energy which has been released, and all the deductions which a nuclear physicist can make from the official statements, are included in the book. It can be recommended to any scientist as a competent survey, carefully written by an expert in this field. The Pilot Press turn out a pleasing volume and are welcomed into this sphere of publication. Their production would have been improved, however, if they had conformed to the standard practice of printing mathematical symbols in italics.

W. COCHRANE.

The Evolution of Modern Physics, by C. T. CHASE. Pp. ix + 203. (New York: D. van Nostrand Co. Inc., 1947.) 14*s*.

There is always much pleasure to be derived from accounts of the many striking incidents and events in the history of physics. Galileo's work on falling bodies ; Fresnel's triumph over the academic doubters ; Röntgen's discovery of x-rays ; Becquerel's discovery of radioactivity—these, and many others, never fail to delight. If the writer manages, at the same time, to show the use of the scientific method on these occasions, it will benefit the general reader, who is patently aware of the achievements of science, but less familiar with the discipline of thought which science requires.

Professor Chase's book contains a well-written and connected account of the history of physics, with a reasonable emphasis on the methods and thought involved, and with practically no mathematics. Some points dealing with the atom and nucleus are rather concentrated and one looks in vain for the apt analogy that would help over an awkward fence.

There are no diagrams or sketches in the book but there are a dozen plates, mainly of complicated pieces of apparatus not referred to in the text. The book is, however, a useful one, since it maintains a high standard. It will help the reputation of physics and will lead people to an understanding of the physicist's work. W. COCHRANE.

The Electron Microscope, by V. E. COSSLETT. Pp. viii + 128. (London: Sigma Books Ltd., 1947.) 7*s*. 6*d*.

The electron microscope is hardly more than 15 years old, yet half a dozen books have been written on it. These books range from the exhaustive monograph to the short introduction for scientists who have no intention to specialize in the subject. The author of the present book now brings principles and uses of the electron microscope " before that wide and increasing public which has a lively interest in the advance of science ".

The content of the book may be divided into three parts of roughly 40 pages each: (1) General introduction, (2) Principles and designs of electron microscopes, (3) Present use and future possibilities of electron microscopes.

The treatment starts in a very general way with some principles of vision, with refraction, prisms and lenses; then it proceeds to the eye, to telescopes and microscopes. In rather quick succession follow explanations on the electron, its refraction and deflection, and on general features of electron lenses. Two excellent chapters on electron microscopes contain details of magnetic and electrostatic types, of shadow- and of scanning-microscopes. Then follows a brief survey of electron microscopical methods and preparation techniques for the investigation of powders, fibres, bacteria, viruses etc. Sectioning and straining methods are discussed and replica methods for the investigation of surfaces are described. A concluding chapter contains information on limitations and possibilities of the electron microscope. These points are of great interest, but perhaps the author has here exceeded the frame of a booklet which is intended for a generally interested public.

The little volume is well illustrated by 40 figures in the text and 12 plates containing a selection of beautiful electron microscopical photographs. The name of the author—who has also written a standard monograph on Electron Optics and who is now in charge of electron microscopic developments in the Cavendish Laboratory in Cambridge—vouches for a masterly treatment of the subject. The book should be warmly recommended to everybody who is interested in electron microscopy. O. KLEMPERER.

A University Text-Book of Physics, Vol. I: *Properties of Matter*, by POYNTING and THOMSON and TODD. Pp. 278. (London: Charles Griffin and Co., Ltd., 1947.) 20*s*.

Of the famous text-book by Poynting and Thomson, the most famous volume was the first, on properties of matter. It is appropriate, therefore, that the first volume to appear of the revision now being undertaken should be that on properties of matter. The original work is so well planned, and so good in its treatment, that one feels reluctant to see too much revision; if an author desires to treat the whole subject differently, it is better that he should write a completely new book. Fortunately, Prof. Todd, the reviser of this volume, and an old student of *both* senior authors, has treated the original text with due respect, and has revised only to the extent that was called for by the developments in the interval since first publication. His chief additions are in the chapters on gravitation, surface tension, and viscosity, and he has certainly made the book a suitable introduction for present-day students. Like others of Griffin's post-war publications, the book is printed on thinner paper than they used to adopt, and the result is to make a pleasanter book to handle. J. H. A.

Frontiers of Astronomy, by DAVID T. EVANS. Pp. 175. (London: Sigma Introductions to Science, 1946.) 6*s*. net.

This is a very good little sketch of the present state of astronomy, giving concise information concerning its chief problems, methods and results. The book is written in a racy, conversational style and is illustrated by a number of diagrams in keeping therewith. A remarkable amount of knowledge is compressed into a small space without introducing obscurity, and although there are, unfortunately, a rather large number of slips, they are mostly of slight importance, and the reader can be satisfied that the impression he gets is an essentially true one. As present prices go the book is strikingly good value for the money. H. D.